

FIGURE 1

GGGGCTTCGGCGCCAGCGGCCAGCGCTAGTCGGTCTGGTAAGGATTTACAAAAGGTGCAGGTATG
AGCAGGTCTGAAGACTAACATTTTGTGAAGTTGTAAAACAGAAAACCTGTTAGAAATGTGGTGGT
TTCAGCAAGGCCTCAGTTTCCTTCCTTCAGCCCTTGTAATTTGGACATCTGCTGCTTTTCATATTT
TCATACATTACTGCAGTAACACTCCACCATATAGACCCGGCTTTACCTTATATCAGTGACACTGG
TACAGTAGCTCCAGAAAAATGCTTATTTGGGGCAATGCTAAATATTGCGGCAGTTTTATGCATTG
CTACCATTTATGTTTCGTTATAAGCAAGTTCATGCTCTGAGTCTGAAGAGAACGTTATCATCAAA
TTAAACAAGGCTGGCCTTGCTACTTGAATACTGAGTTGTTTAGGACTTTCTATTGTGGCAAACCTT
CCAGAAAACAACCCCTTTTGTGCTGCACATGTAAGTGGAGCTGTGCTTACCTTTGGTATGGGCTCAT
TATATATGTTTGTTCAGACCATCCTTTCTACCAAATGCAGCCCAAAATCCATGGCAAACAAGTC
TTCTGGATCAGACTGTTGTTGGTTATCTGGTGTGGAGTAAGTGCACCTTAGCATGCTGACTTGCTC
ATCAGTTTTGCACAGTGGCAATTTTGGGACTGATTTAGAACAGAACTCCATTGGAACCCCGAGG
ACAAAGGTTATGTGCTTCACATGATCACTACTGCAGCAGAATGGTCTATGTCAATTTTCCTTCTTT
GGTTTTTTCCTGACTTACATTCGTGATTTTCAGAAAATTTCTTTACGGGTGGAAGCCAATTTACA
TGGATTAACCCCTCTATGACACTGCACCTTGCCCTATTAACAATGAACGAACACGGCTACTTTCCA
GAGATATTGATGAAAGGATAAAATATTTCTGTAATGATTATGATTCTCAGGGATTGGGGAAAGG
TTCACAGAAGTTGCTTATTTCTCTCTGAAATTTCAACCACTTAATCAAGGCTGACAGTAACACT
GATGAATGCTGATAATCAGGAAACATGAAAGAAGCCATTTGATAGATTATTCTAAAGGATATCAT
CAAGAAGACTATTA AAAACACCTATGCCTATACTTTTTTATCTCAGAAAATAAAGTCAAAAGACT
ATG

FIGURE 2

<subunit 1 of 1, 266 aa, 1 stop

<MW: 29766, pI: 8.39, NX(S/T): 0

MWWFQQGLSFLPSALVIWTSAAAFIFSITAVTLHHIDPALPYISDTGTVAPEKCLFGAMLNIAAV
LCIATIIYVRYKQVHALSPEENVIIKLNKAGLVLGILSCLGLSIVANFQKTTLFAAHVSGAVLTFG
MGSLYMFVQTILSYQMOPKIHGKQVFWIRLLLVIWCGVSALSMLTCSSVLHSGNFGTDLEQKLHW
NPEDKGYVLHMITTAAEWSMSFSFFGFFLTYYIRDFQKISLRVEANLHGLTLYDTAPCPINNERTR
LLSRDI

Important features:

Type II transmembrane domain:

amino acids 13-33

Other Transmembrane domains:

amino acids 54-73, 94-113, 160-180, 122-141

N-myristoylation sites.

amino acids 57-63, 95-101, 99-105, 124-130, 183-189

FIGURE 3

CGGACGCGTGGGCGGACGCGTGGGGGAGAGCCGCGAGTCCCGGCTGCAGCACCTGGGAGAAGGCAGACC
GTGTGAGGGGGCTGTGGCCCCAGCGTGTGTGGCCTCGGGGAGTGGGAAGTGGAGGCAGGAGCCTTC
CTTACACTTCGCCATGAGTTTCCTCATCGACTCCAGCATCATGATTACCTCCCAGATACTATTTTTTG
GATTTGGGTGGCTTTTCTTCATGCGCCAATTGTTTAAAGACTATGAGATACGTGAGTATGTTGTACAG
GTGATCTTCTCCGTGACGTTTGCATTTTCTTGACCATGTTTGAGCTCATCATCTTTGAAATCTTAGG
AGTATTGAATAGCAGCTCCCGTTATTTTCACTGGAAAATGAACCTGTGTGAATTCTGCTGATCCTGG
TTTTTCATGGTGCCTTTTTACATTGGCTATTTTTATTGTGAGCAATATCCGACTACTGCATAAACAACGA
CTGCTTTTTTCTGTCTCTTATGGCTGACCTTTATGTATTTCTTCTGGAACTAGGAGATCCCTTTCC
CATTCTCAGCCCAAAACATGGGATCTTATCCATAGAACAGCTCATCAGCCGGGTGGTGTGATTGGAG
TGACTCTCATGGCTCTTCTTTCTGGATTGGTGTGTCAACTGCCCATAACCTTACATGTCTTACTTC
CTCAGGAATGTGACTGACACGGATATTCTAGCCCTGGAACGGCGACTGCTGCAAACCATGGATATGAT
CATAAGCAAAAAGAAAAGGATGGCAATGGCACGGAGAACAATGTTCCAGAAGGGGGAAGTGATAACA
AACCATCAGGTTTCTGGGGAATGATAAAAAGTGTACCACCTCAGCATCAGGAAGTGAAAATCTTACT
CTTATTCAACAGGAAGTGGATGCTTTGGAAGAATTAAGCAGGCAGCTTTTTCTGGAAACAGCTGATCT
ATATGCTACCAAGGAGAGAATAGAATACTCCAAAACCTTCAAGGGGAAATATTTTAATTTTCTTGGTT
ACTTTTTCTCTATTTACTGTGTTTGAAAATTTTCATGGCTACCATCAATATTGTTTTTGATCGAGTT
GGGAAAACGGATCCTGTCAAGAGGCATTGAGATCACTGTGAATTATCTGGGAATCCAATTGATGT
GAAGTTTGGTCCCAACACATTTCTTCATTCTTGTGGAATAATCATCGTCACATCCATCAGAGGAT
TGCTGATCACTCTTACCAAGTTCTTTTATGCCATCTCTAGCAGTAAGTCCTCCAATGTCATTGTCTGT
CTATTAGCACAGATAATGGGCATGTACTTTGTCTCCTCTGTGCTGATCCGAATGAGTATGCCTTT
AGAATACCGCACCATTAATCACTGAAGTCCTTGGAAGTGCAGTTCAACTTCTATCACCGTTGGTTTG
ATGTGATCTTCTGGTCAGCGCTCTCTCTAGCATACTCTTCTCTATTTGGCTCACAAACAGGCACCA
GAGAAGCAAATGGCACCTTGAACTTAAGCCTACTACAGACTGTTAGAGGCCAGTGGTTTCAAATTTA
GATATAAGAGGGGGGAAAAATGGAACAGGGCCTGACATTTTATAAACAACAAAATGCTATGGTAGC
ATTTTTCACCTTCATAGCATACTCCTTCCCGTCAGGTGATACTATGACCATGAGTAGCATCAGCCAG
AACATGAGAGGGGAGAACTAACTCAAGACAATACTCAGCAGAGAGCATCCCGTGTGGATATGAGGCTGG
TGTAAGGCGGAGAGGAGCAAGAACTAAAGGTGAAAAATACACTGGAACCTCTGGGGCAAGACATGT
CTATGGTAGCTGAGCCAAACACGTAGGATTTCCGTTTTAAGGTTTACATGGAAAAGGTTATAGCTTTG
CCTTGAGATTGACTCATTAATCAGAGACTGTAACAAAAAAAAAAAAAAAAAAGGGCGGCCGCG
ACTCTAGAGTCGACCTGCAGAAGCTTGCCGCCATGGCCCAACTTGTTATTGCAGCTTATAATG

FIGURE 4

MSFLIDSSIMITSQILFFGFWLFFMRQLFKDYEIRQYVVQVIFSVTFAFSCTMFELIIFEILGV
LNSSRYFHWKMNL CVILLILVFMVPFYIGYFIVSNIRLLHKQRLLFSCLLWLT FMYFFWKL GDP
FPILSPKHGILSIEQLISRVGVIGVTLMALLSGFGAVNCPYTYMSYFLRNVTDTDILALERRLLQ
TMDMIISKKKRMAMARRTMFQKGEVHNKPSGFWGMIKSVTTSASGSENLT LIQQEVDAL EELSRQ
LFLETADLYATKERIEYSKTFKGKYFNFLGYFFSIYCVWKIFMATINIVFDRVGKTD PVTRGIEI
TVNYLGIQFDVKFWSQHISFILVGIIIVTSIRGLLITLTKFFYAISSSKSSNVIVLLLAQIMGY
FVSSVLLIRMSMPLEYRTIITEVLGELQFNFYHRWFDVIFLVSALSSILFLYLAHKQAPEKQMAP

Important features:

Signal peptide:

amino acids 1-23

Potential transmembrane domains:

amino acids 37-55, 81-102, 150-168, 288-311, 338-356, 375-398,
425-444

N-glycosylation sites.

amino acids 67-70, 180-183 and 243-246

Eukaryotic cobalamin-binding proteins

amino acids 151-160

FIGURE 5

AGCAGGGAAATCCGGATGTCTCGGTTATGAAGTGGAGCAGTGAGTGTGAGCCTCAACATAGTTCC
 AGAACTCTCCATCCGGACTAGTTATTGAGCATCTGCCTCTCATATCACCAGTGGCCATCTGAGGT
 GTTTCCTGGCTCTGAAGGGGTAGGCACGATGCCAGGTGCTTCAGCCTGGTGTGCTTCTCACT
 TCCATCTGGACCACGAGGCTCCTGGTCCAAGGCTCTTTGCGTGCAGAAGAGCTTCCATCCAGGT
 GTCATGCAGATTATGGGGATCACCCCTTGTGAGCAAAAAGGCGAACCAGCAGCTGAATTTACAG
 AAGCTAAGGAGGCTGTAGGCTGCTGGGACTAAGTTTGGCCGGCAAGGACCAAGTTGAAACAGCC
 TTGAAAGCTAGCTTTGAAACTTGCAGCTATGGCTGGGTGGAGATGGATTTCGTGGTTCATCTCTAG
 GATTAGCCCAACCCCAAGTGTGGGAAAAATGGGGTGGGTGTCCTGATTTGGAAGGTTCAGTGA
 GCCGACAGTTTGCAGCCTATTGTTACAACCTCATCTGATACTTGGACTAACTCGTGCATTCCAGAA
 ATTATCACCACCAAGATCCCATATTCAACACTCAACTGCAACACAAACAACAGAATTTATGT
 CAGTGACAGTACCTACTCGGTGGCATCCCTTACTCTACAATACCTGCCCCCTACTACTCCTC
 CTGCTCCAGCTTCCACTTCTATTCCACGGAGAAAAAATTGATTTGTGTACAGAAGTTTTTATG
 GAACTAGCACCATGTCTACAGAACTGAACCATTTGTTGAAAATAAAGCAGCATTCAAGAATGA
 AGCTGCTGGGTTTGGAGGTTCACCGCTCTGCTAGTGCTTGCTCTCCTCTTCTTTGGTGTG
 CAGCTGGTCTTGGATTTTGTATGTCAAAGGTATGTGAAGGCCTTCCCTTTTACAAACAAGAA
 CAGCAGAAGGAAATGATCGAAACCAAGTAGTAAAGGAGGAGAAGGCCAATGATAGCAACCCATA
 TGAGGAATCAAAGAACTGATAAAACCCAGAAGAGTCCAAGAGTCCAAGCAAACTACCGTGC
 GATGCCCTGGAAGCTGAAGTTTAGATGAGACAGAAATGAGGAGACACACCTGAGGCTGGTTCTTT
 CATGCTCCTTACCTGCCCCAGCTGGGGAATCAAAGGGCCAAAGAACCAAGAAGAAAGTCCA
 CCCTTGGTTTCTAACTGGAATCAGCTCAGGACTGCCATTGGACTATGGAGTGCACCAAGAGAAT
 GCCCTTCTCCTTATTGTAACCTGTCTGGATCCTATCCTCCTACCTCCAAAGCTTCCCACGGCCT
 TTCTAGCCTGGCTATGTCTTAATAATATCCCACTGGGAGAAAGGAGTTTTCGAAAGTGAAGGAC
 CTAAACATCTCATCAGTATCCAGTGGTAAAAAGGCCTCCTGGCTGTCTGAGGCTAGGTGGGTG
 AAAGCCAAGGAGTCACTGAGACCAAGGCTTCTCTACTGATTCCGCAGCTCAGACCCCTTCTTCA
 GCTCTGAAAGAGAAACACGTATCCACCTGACATGTCCTTCTGAGCCCGGTAAGAGCAAAAGAAT
 GGCAGAAAAGTTTAGCCCTGAAAGCCATGGAGATTCTCATAACTTGAGACCTAATCTCTGTAAA
 GCTAAAATAAAGAAATAGAACAAGGCTGAGGATACGACAGTACACTGTGAGCAGGGACTGTAAAC
 ACAGACAGGGTCAAAGTGTTTTCTCTGAACACATTGAGTTGGAATCACTGTTTGAACACACACA
 CTTACTTTTTCTGGTCTCTACCACTGCTGATATTTCTCTAGGAAATATACTTTTACAAGTAACA
 AAAATAAAACTCTTATAAATTTCTATTTTTATCTGAGTTACAGAAATGATTACTAAGGAAGATT
 ACTCAGTAATTTGTTTAAAAAGTAATAAAATTCACAAACATTTGCTGAATAGCTACTATATGTC
 AAGTGCTGTGCAAGGTATTACACTCTGTAATTGAATATTATTCCTCAAAAAATTGCACATAGTAG
 AACGCTATCTGGGAAGCTATTTTTTTCAGTTTTGATATTTCTAGCTTATCTACTTCCAACTAAT
 TTTTATTTTGTGAGACTAATCTTATTCATTTTCTCTAATATGGCAACCATTATAACCTTAATT
 TATTATTAACATACCTAAGAAGTACATTGTTACCTCTATATACCAAGCACATTTTAAAGTGCC
 ATTAACAAATGTATCACTAGCCCTCCTTTTTCCAACAAGAAGGACTGAGAGATGCAGAAATATT
 TGTGACAAAAAATTAAAGCATTTAGAAACTT

FIGURE 6

MARCFSLVLLLTISIWTTRLLVQGSRLAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEACRLLG
LSLAGKDQVETALKASFETCSYGWVGDFVVISRISPNPKCGKNGVGVLIWKVPVSRQFAAYCYN
SSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPAPTTTPAPASTSIPR
RKKLICVTEVEMETSTMSTETEPFVENKAAFKNEAAGFGGVPTALLVLALLFFGAAAGLGFCYVK
RYVKAFPFNTKNQOKEMIETKVVKEEKANDSNPNEESKKTDKNPESKSPSKTTVRCLEAEV

Signal sequence:

amino acids 1-16

Transmembrane domain:

amino acids 235-254

N-glycosylation site.

amino acids 53-57, 130-134, 289-293

Casein kinase II phosphorylation site.

amino acids 145-149, 214-218

Tyrosine kinase phosphorylation site.

amino acids 79-88

N-myristoylation site.

amino acids 23-29, 65-71, 234-240, 235-239, 249-255, 253-259

FIGURE 7

CGCCGCGCTCCCGCACCCGCGGCCCGCCACCGCGCGCTCCCGCATCTGCACCCGAGCCCGGC
GGCCTCCCGGCGGGAGCGAGCAGATCCAGTCCGGCCCGCAGCGCAACTCGGTCCAGTCCGGGCGG
CGGCTGCGGGCGCAGAGCGGAGATGCGAGCGGCTTGGGGCCACCCTGCTGTGCCTGCTGTGCGG
CGGCGGTCCCCACGGCCCCCGCGCCGCTCCGACGGCGACCTCGGCTCCAGTCAAGCCCGGCCG
GCTCTCAGCTACCCGCGAGGAGGAGGCCACCCTCAATGAGATGTTCCGCGAGGTTGAGGAAGTAT
GGAGGACACGAGCACAATTCGCGAGCGCGGTGGAAGAGATGGAGGCAGAAGAAGCTGCTGCTA
AAGCATCATCAGAAGTGAACCTGGCAAACCTTACCTCCAGCTATCACAATGAGACCAACACAGAC
ACGAAGGTTGAAATAATACCATCCATGTGCACCGAGAAATTCACAAGATAACCAACAACCGAC
TGGACAATGGTCTTTTCAGAGACAGTTATCACATCTGTGGGAGACGAAGAAGGCAGAAGGAGCC
ACGAGTGCATCATCGACGAGGACTGTGGGCCAGCATGTACTGCCAGTTTGGCAGCTTCCAGTAC
ACCTGCCAGCCATGCCGGGGCCAGAGGATGCTCTGCACCCGGGACAGTGAAGTGTGTGGAGACCA
GCTGTGTGTCTGGGGTCACTGCACCAAAATGGCCACCAGGGGCAGCAATGGGACCATCTGTGACA
ACCAGAGGGACTGCCAGCCGGGGCTGTGCTGTGCCTTCCAGAGAGGCTGCTGTTCCCTGTGTGC
ACACCCCTGCCCGTGGAGGGCGAGCTTTGCCATGACCCGCCAGCCGGCTTCTGGACCTCATCAC
CTGGGAGCTAGAGCCTGATGGAGCCTTGGACCGATGCCCTTGTGCCAGTGGCCTCCTCTGCCAGC
CCCACAGCCACAGCCTGGTGTATGTGTGCAAGCCGACCTTCGTGGGGAGCCGTGACCAAGATGGG
GAGATCCTGCTGCCAGAGAGGTCCCCGATGAGTATGAAGTTGGCAGCTTCATGGAGGAGGTGCG
CCAGGAGCTGGAGGACCTGGAGAGGAGCCTGACTGAAGAGATGGCGCTGGGGGAGCCTGCGGCTG
CCGCCGCTGCACTGCTGGGAGGGGAAGAGATTAGATCTGGACCAGGCTGTGGGTAGATGTGCAA
TAGAAATAGCTAATTTATTTCCCCAGGTGTGTGCTTTAGGCGTGGGCTGACCAGGCTTCTTCCTA
CATCTTCTTCCAGTAAGTTTCCCTCTGGCTTGACAGCATGAGGTGTGTGCATTGTTCAGT
CCCCAGGCTGTTCTCCAGGCTTACAGTCTGGTGTGGGAGAGTCAAGGAGGTTAAACTGCA
GGAGCAGTTTGGCAACCCCTGTCCAGATTATGGCTGCTTTGCCTCTACCAGTTGGCAGACAGCCG
TTTGTCTACATGGCTTTGATAATTGTTTGGGGGAGGAGATGGAACAATGTGGAGTCTCCCTC
TGATTGGTTTTGGGGAATGTGGAGAAGAGTGCCCTGCTTTGCAACATCAACCTGGCAAAATG
CAACAAATGAATTTCCACGCAGTTCTTTCCATGGGCATAGGTAAGCTGTGCCTTCAGCTGTTGC
AGATGAAATGTTCTGTTACCCCTGCATTACATGTGTTTATTCATCCAGCAGTGTGCTCAGCTCC
TACCTCTGTGCCAGGGCAGCATTTCATATCCAAGATCAATTCCTCTCTCAGCACAGCCTGGGG
AGGGGGTCATTGTTCTCCTCGTCCATCAGGGATCTCAGAGGCTCAGAGACTGCAAGCTGCTTGCC
CAAGTCACACAGCTAGTGAAGACCAGAGCAGTTTCATCTGGTTGTGACTCTAAGCTCAGTGTCT
CTCCACTACCCACACAGCCTTGGTGCCACCAAAAGTGCTCCCCAAAAGGAAGGAGAATGGGAT
TTTTCTTGAGGCATGCACATCTGGAATTAAGGTCAAACATATCTCACATCCCTCTAAAAGTAAA
CTACTGTTAGGAACAGCAGTGTCTCACAGTGTGGGGCAGCCGTCCTTCTAATGAAGACAATGAT
ATTGACACTGTCCCTCTTTGGCAGTTGCATTAGTAACCTTGAAGGTATATGACTGAGCGTAGCA
TACAGGTTAACCTGCAGAAACAGTACTTAGGTAATTGTAGGGCGAGGATTATAAATGAATTTGC
AAAATCACTTAGCAGCAACTGAAGACAATTATCAACCACGTGGAGAAAATCAAACCGAGCAGGGC
TGTGTGAAACATGGTTGTAATATGCGACTGCGAACAAGTGAAGTCTACGCCACTCCACAAATGATG
TTTTCAGGTGTGATGAGTGTGTGCCACCATGTATTATCCAGAGTCTTAAAGTTTAAAGTTGCA
CATGATTGTATAAGCATGCTTTCTTTGAGTTTTAAATTATGTATAAACATAAGTTGCATTAGAA
ATCAAGCATAAATCACTTCACTGCAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 8

MQRLGATLLCLLLAAVPTAPAPAPTATSAPVKPGPALSYQEEATLNEMFREVEELMEDTQHKL
RSVEEMEAEAAAAKASSEVNLANLPPSYHNETNTDTKVGNNTIHVHREIHKITNNQTGQMVFSE
TVITSVGDEEGRRSHECIIDEDCGPSMYCQFASFQYTCQPCRQRMCLTRDSECCGDQLCVWGHC
TKMATRGSNGTICDNQRDCQPGLCCAFQRGLLFPVCTPLPVEGELCHDPASRLLDLITWELEPDG
ALDRPCASGLLCQPHSHSLVYVCKPTFVGSRDQDGEILLPREVPDEYEVGSFMEEVRQELEDLE
RSLTEEMALGEPAAAAAALLGGEEI

Signal sequence:

amino acids 1-19

N-glycosylation site.

amino acids 96-100, 106-110, 121-125, 204-208

Casein kinase II phosphorylation site.

amino acids 46-50, 67-71, 98-102, 135-139, 206-210, 312-316,
327-331

N-myristoylation site.

amino acids 202-208, 217-223

Amidation site.

amino acids 140-144

FIGURE 9

CGGACGCGTGGGCGGACGCGTGGGGCTGTGAGAAAGTGCCAATAAATACATCATGCAACCCAC
GGCCACCTTGTGAACCTCGTGCCAGGGCTGATGTGCGTCTTCCAGGGCTACTCATCCAAAG
GCCTAATCCAACGTTCTGTCTTCAATCTGCAAATCTATGGGGTCTGGGGCTCTTCTGGACCCCT
AACTGGGTACTGGCCCTGGGCCAATGCGTCTCGCTGGAGCCTTTCCTCCTTCTACTGGGCCTT
CCACAAGCCCCAGGACATCCCTACCTTCCCCTTAATCTCTGCCTTCATCCGCACACTCCGTACC
ACACTGGGTCAATTGGCATTGAGCCCTCATCTGACCCTTGTGCAGATAGCCGGGTATCTTG
GAGTATATTGACCACAAGCTCAGAGGAGTGCAGAACCCTGTAGCCGCTGCATCATGTGCTGTTT
CAAGTGCTGCCTCTGGTGTCTGGAAAAATTTATCAAGTTCCTAAACCGCAATGCATACATCATGA
TCGCCATCTACGGGAAGAATTTCTGTGTCTCAGCCAAAAATGCGTTCATGCTACTCATGCGAAAC
ATTGTCAGGGTGGTCGTCTGGACAAAGTCACAGACCTGCTGCTGTTCTTTGGGAAGCTGCTGGT
GGTCGGAGGCGTGGGGGTCTGTCTTCTTTTTTCTCCGGTCGCATCCCGGGGCTGGGTAAAG
ACTTTAAGAGCCCCACCTCAACTATTACTGGCTGCCCATCATGACCTCCATCCTGGGGGCTAT
GTCATCGCCAGCGGCTTCTTCAGCGTTTTTCGGCATGTGTGGACACGCTCTTCTCTGCTTCCT
GGAAGACCTGGAGCGGAACAACGGCTCCCTGGACCGCCCTACTACATGTCCAAGAGCCTTCTAA
AGATTCTGGGCAAGAAGAACGAGGCGCCCCCGGACAACAAGAAGAGGAAGAAGTGACAGCTCCGG
CCCTGATCCAGGACTGCACCCACCCACCGTCCAGCCATCCAACCTCACTTCGCCTTACAGGT
CTCCATTTTGTGGTAAAAAAGGTTTTAGGCCAGGCGCGTGGCTCACGCCTGTAATCCAACACT
TTGAGAGGCTGAGGCGGGCGGATCACCTGAGTCAGGAGTTCGAGACCAGCCTGGCCAACATGGTG
AAACCTCCGTCTCTATTAATAAATACAAAAATTAGCCGAGAGTGGTGGCATGCACCTGTCATCCCA
GCTACTCGGGAGGCTGAGGCAGGAGAATCGCTTGAACCCGGGAGGCAGAGGTGTCAGTGAGCCGA
GATCGCGCCACTGCACTCCAACCTGGGTGACAGACTCTGTCTCCAAAACAAAACAAACAAA
AAGATTTTATTAAAGATATTTTGTTAACTC

FIGURE 10

RTGRTRGGCEKVPINTSCNPTAHLVNSSCPGLMCVFQGYSSKGLIQRSVFNLQIYGVLGLEFWTL
NWVLALGQCVLAFASFYWAFHKPDIPTFPLISAFIRTLRYHTGSLAFGALILTLVQIARVIL
EYIDHKLRGVQNPVARCIMCCFKCCLWCLEKFIKFLNRNAYIMIAIYGKNFCVSAKNAFMLLMRN
IVRVVLDKVTDLLLFFGKLLVVGGVLSFFFFSGRIPGLGKDFKSPHLNYYWLPIMTSILGAY
VIASGFFSVFGMCVDTLFLCFLEDLERNNGSLDRPYMSKSLKILGKKNEAPPDNKKRKK

Important features:

Transmembrane domains:

amino acids 57-80 (type II), 110-126, 215-231, 254-274

N-glycosylation sites.

amino acids 16-20, 27-31, 289-293

Hypothetical YBR002c family proteins.

amino acids 276-288

Ammonium transporters proteins.

amino acids 204-231

N-myristoylation sites.

amino acids 60-66, 78-84

Amidation site.

amino acids 306-310

FIGURE 11

GCCCCGCGCCCGGCGCCGGGCGCCCCGAAGCCGGGAGCCACCGCCATGGGGGCCTGCCTGGGAGCCTGC
TCCCTGCTCAGCTGCGCGTCCCTGCCTCTGCGGCTCTGCCCCCTGCATCCTGTGCAGCTGCTGCCCCGC
CAGCCGCAACTCCACCGTGAGCCGCCTCATCTTCACGTTCTTCTCTTCTGGGGGTGCTGGTGTCCA
TCATTATGCTGAGCCCGGGCGTGAGAGTCAGCTCTACAAGCTGCCCTGGGTGTGTGAGGAGGGGGCC
GGGATCCCCACCGTCCCTGCAGGGCCACATCGACTGTGGCTCCCTGCTTGGCTACCGCGCTGTCTACCG
CATGTGCTTCGCCACGGCGGCTTCTTCTTCTTTTTCACCTGCTCATGCTCTGCGTGAGCAGCA
GCCCCGACCCCGGCTGCCATCCAGAATGGGTTTGGTTCTTTAAGTTCCTGATCCTGGTGGGCCTC
ACCGTGGGTGCCTTCTACATCCCTGACGGCTCCTTCACCAACATCTGGTTCTACTTCGGCGTCGTGGG
CTCCTTCTCTTTCATCCTCATCCAGCTGGTGCTGCTCATCGACTTTGCGCACTCCTGGAACAGCGGT
GGCTGGGCAAGGCCGAGGAGTGCATTCCCGTGCCCTGGTACGCAGGCCTCTTCTTCTTCACTCTCCTC
TTCTACTTGCTGTGATCGCGGCGTGCGCTGATGTTTATGTACTACTGAGCCAGCGGCTGCCA
CGAGGGCAAGGTCTTTCATCAGCTCAACCTCACCTTCTGTGTCTGCGTGTCCATCGCTGCTGCTGCTG
CCAAGGTCCAGGACGCCAGCCCAACTCGGGTCTGCTGCAGGCCTCGGTATCACCTCTACACCATG
TTTGTACCTGGTCAGCCCTATCCAGTATCCCTGAACAGAAATGCAACCCCATTTGCCAACCCAGCT
GGGCAACGAGACAGTTGTGGCAGGCCCCGAGGGCTATGAGACCCAGTGGTGGGATGCCCCGAGCATTG
TGGGCTCATCATCTTCTCCTGTGCACCTCTTTCATCAGTCTGCGCTCCTCAGACCACGGCAGGTG
AACAGCCTGATGCAGACCGAGGAGTGGCCACCTATGCTAGACGCCACACAGCAGCAGCAGCAGGT
GGCAGCCTGTGAGGGCCGGGCTTTGACAACGAGCAGGACGGCGTCACCTACAGCTACTCCTTCTTCC
ACTTCTGCCCTGGTGTGGCCTCACTGCACGTCATGATGACGCTCACCAACTGGTACAAGCCCGGTGAG
ACCCGGAAGATGATCAGCACGTGGACCGCGTGTGGGTGAAGATCTGTGCCAGCTGGGCAGGGCTGCT
CCTCTACCTGTGGACCCTGGTAGCCCCACTCCTCCTGCGCAACCGCGACTTCAGCTTGAGGCAGCCTCA
CAGCCTGCCATCTGGTGCCTCCTGCCACCTGGTGCCTCTCGGCTCGGTGACAGCCAACTGCCCCCTC
CCCACACCAATCAGCCAGGCTGAGCCCCCACCCTGCCCCAGCTCCAGGACCTGCCCTGAGCCGGGC
CTTCTAGTCGTAGTGCTTTCAGGGTCCGAGGAGCATCAGGCTCCTGCAGAGCCCCATCCCCCGCCAC
ACCCACACGGTGGAGCTGCCTCTTCTTCCCTCCTCCTGTTGCCATACTCAGCATCTCGGATGAA
AGGGCTCCCTTGTCTCAGGCTCCACGGGAGCGGGGCTGCTGGAGAGAGCGGGGAATCCACCACAG
TGGGGCATCCGGCACTGAAGCCCTGGTGTCTTCTGGTACGTCCTCCAGGGGACCCTGCCCCCTTCTG
GACTTCGTGCCCTTACTGAGTCTCTAAGACTTTTCTAATAAACCAAGCCAGTGCGGTGTAACAAAAA

FIGURE 12

MGACLGACSLSCASCLCGSAPCILCSCCPASRNSTVSRLIFTFFLFLGVLSIIMLSPGVESQL
YKLPWVCEEAGIPTVLQGHIDCGSLLGYRAVYRMCFATAAFFFFFFFFTLLMLCVSSSRDPRAAIQ
NGFWFFKFLILVGLTVGAFYIPDGSFTNIWFYFGVVGSFLFILIQLVLLIDFAHSWNQRWLKAE
ECDSRAWYAGLFFFTLLFYLLSIAAVALMFMYYTEPSGCHGKVFISLNLTCVCVSIAAVLPKV
QDAQPNSGLLQASVITLYTMFVTWSALSSIPEQKCNPHLPTQLGNETVVAGPEGYETQWWDAPSI
VGLIIFLLCTLFISLRSSDHRQVNSLMQTEECPPMLDATQQQQQVAACEGRAFDNEQDGVITYSY
SFFHFCLVLASLHVMTLTNWYKPGETRKMISTWTAVWVKICASWAGLLLYLWTLVAPLLLRNRD
FS

Signal sequence:

amino acids 1-20

Transmembrane domains:

amino acids 40-58, 101-116, 134-150, 162-178, 206-223, 240-257,
272-283, 324-340, 391-406, 428-444

FIGURE 13

CGGGCCAGCCTGGGGCGGCCGGCCAGGAACCACCCGTTAAGGTGTCTTCTCTTTAGGGATGGTGA
GGTTGGA AAAAGACTCCTGTAACCCCTCCTCCAGGATGAACCACCTGCCAGAAGACATGGAGAACG
CTCTCACC GGGAGCCAGAGCTCCCATGCTTCTCTGCGCAATATCCATTCCATCAACCCACACAA
CTCATGGCCAGGATTGAGTCCATGAAGGAAGGAAAAGAAAGGCATATCTGATGTCAGGAGGAC
TTTCTGTTTGTGTTGTACCTTTGACCTCTTATTTCGTAACATTACTGTGGATAATAGAGTTAAATG
TGAATGGAGGCATTGAGAACACATTAGAGAAGGAGGTGATGCAGTATGACTACTATTCTTCATAT
TTTGATATATTCTTCTGGCAGTTTTTCGATTTAAAGTGTTAATACTTGCATATGCTGTGTGCAG
ACTGCGCCATTGGTGGGCAATAGCGTTGACAACGGCAGTGACCAGTGCCTTTTTACTAGCAAAAAG
TGATCCTTTTCGAAGCTTTTCTCTCAAGGGGCTTTTGGCTATGTGCTGCCCATCATTTCATTATC
CTTGCCCTGGATTGAGACGTGGTTCCTGGATTTCAAAGTGTTACCTCAAGAAGCAGAAGAAGAAAA
CAGACTCCTGATAGTTTCCAGGATGCTTCAGAGAGGGCAGCACTTATACCTGGTGGTCTTCTGATG
GTCAGTTTTATTCCCCTCCTGAATCCGAAGCAGGATCTGAAGAAGCTGAAGAAAAACAGGACAGT
GAGAAACCACCTTTTAGAACTATGAGTACTACTTTTGTAAATGTGAAAAACCCTCACAGAAAGTC
ATCGAGGCCAAAAAGAGGCAGGCAGTGGAGTCTCCCTGTCGACAGTAAAGTTGAAATGGTGACGTC
CACTGCTGGCTTTATTGAACAGCTAATAAAGATTTATTTATTGTAATACCTCACAAACGTTGTAC
CATATCCATGCACATTTAGTTGCCTGCCTGTGGCTGGTAAGGTAATGTCATGATTCATCCTCTCT
TCAGTGAGACTGAGCCTGATGTGTTAACAAATAGGTGAAGAAAGTCTTGTGCTGTATTCTTAATC
AAAAGACTTAATATATTGAAGTAACATTTTTTAGTAAGCAAGATACCTTTTTATTTCATTCAC
AGAAATGGAATTTTTTTGTTTCATGTCTCAGATTTATTTGTATTTCTTTTTTAACACTCTACATT
TCCCTTGTTTTTTAACTCATGCACATGTGCTCTTGTACAGTTTTAAAAAGTGAATAAAATCTG
ACATGTCAATGTGGCTAGTTTTATTTTTCTGTTTTGCATTATGTGTATGGCCTGAAGTGTTGGA
CTTGCAAAAGGGGAAGAAAGGAATTGCGAATACATGTAAAATGTCACCAGACATTTGTATTATTT
TTATCATGAAATCATGTTTTTCTGATTGTCTGAAATGTTCTAAATACTCTTATTTTGAATGC
ACAAAATGACTTAAACCATTATATCATGTTTCCTTTGCGTTCAGCCAATTTCAATTAAAATGAA
CTAAATTAAAAA

FIGURE 14

MNHLPEDMENALTGSQSSHASLRNIHSINPTQLMARIESYEGREKKGISDVRRFTCLFVTFDLLF
VTLLWIIELNVNGGIENTLEKEVMQYDYSSYFDIFLLAVFRFKVLILAYAVCRLRHWWAIALTT
AVTSAFLLAKVILSKLFSQCAFYVLPPIISFILAWIETWFLDFKVLQPQEAEEENRLLIVQDASER
AALIPGGLSDGQFYSPPESEAGSEEAEKQDSEKPLLEL

Important features of the protein:

Signal peptide:

amino acids 1-20

Transmembrane domains:

amino acids 54-72, 100-118, 130-144, 146-166

N-myristoylation sites.

amino acids 14-20, 78-84, 79-85, 202-208, 217-223

FIGURE 15

ACTCGAACGCAGTTGCTTCGGGACCCAGGACCCCTCGGGCCCCGACCCGCCAGGAAAGACTGAGG
 CCGCGGGCTGCCCGCCCGGCTCCCTGCGCCGCCCGCCTCCCGGGACAGAAGATGTGCTCCAG
 GGTCCCTCTGCTGCTGCCGCTGCTCCTGCTACTGGCCCTGGGGCTGGGGTGCCAGGGCTGCCCAT
 CCGGCTGCCAGTGCCAGCCAGCCACAGACAGTCTTCTGCACTGCCCGCCAGGGGACCACGGTGCCC
 CGAGACGTGCCACCCGACACGGTGGGGCTGTACGTCTTTGAGAACGGCATCACCATGCTCGACGC
 AGGCAGCTTTGCCGGCTGCCGGGCTGCAGCTCCTGGACCTGTACAGAACCCAGATGCCAGCC
 TGCCAGCGGGTCTTCCAGCCACTCGCCAACTCAGCAACCTGGACCTGACGGCCAAACAGGCTG
 CATGAAATCACCAATGAGACCTTCCGTGGCCTGCGGCGCCTCGAGCGCCTCTACCTGGGCAAGAA
 CCGCATCCGCCACATCCAGCCTGGTGCCCTTCGACACGCTCGACCGCCTCCTGGAGCTCAAGCTGC
 AGGACAACGAGCTGCGGGCACTGCCCCGCTGCGCCTGCCCCGCTGCTGCTGCTGGACCTCAGC
 CACAACAGCCTCCTGGCCCTGGAGCCCGGCATCCTGGACACTGCCAACGTGGAGGCGCTGCGGCT
 GGCTGGTCTGGGGCTGCAGCAGCTGGACGAGGGGCTCTTCAGCCGCTTGCGCAACCTCCACGACC
 TGGATGTGTCCGACAACAGCTGGAGCGAGTGCCACCTGTGATCCGAGGCCTCCGGGGCTGACG
 CGCCTGCGGCTGGCCGGCAACCCCGCATTCGCCAGCTGCGGGCCGAGGACCTGGCCGGCCTGGC
 TGCCCTGCAGGAGCTGGATGTGAGCAACCTAAGCCTGCAGGCCCTGCCTGGCGACCTCTCGGGCC
 TCTTCCCCCGCTGCGGCTGCTGGCAGCTGCCCCGAACCCCTTCAACTGCGTGTCCCCCTGAGC
 TGGTTTGGCCCTGGGTGCGGAGAGCCACGTCACTGGCCAGCCCTGAGGAGACGCGCTGCCA
 CTTCGCCCCCAAGAACGCTGGCCGGCTGCTCCTGGAGCTTGAATACGCCGACTTTGGCTGCCAG
 CCACCACCACACAGCCACAGTGCCACCAACGAGGCCCGTGGTGCGGGAGCCCACAGCCTTGTCT
 TCTAGCTTGGCTCCTACCTGGCTTAGCCCCACAGCGCCGGCCACTGAGGCCCCAGCCCGCCCTC
 CACTGCCCCACCGACTGTAGGGCCTGTCCCCAGCCCCAGGACTGCCACCGTCCACCTGCCTCA
 ATGGGGGCACATGCCACCTGGGACACGGCACCACTGGCGTGCTTGTGCCCCGAAGGCTTCACG
 GGCTGTACTGTGAGAGCCAGATGGGGCAGGGGACACGGCCAGCCCTACACAGTCACGCCGAG
 GCCACCACGGTCCCTGACCTGGGCATCGAGCCGGTGAGCCCCACCTCCCTGCGCGTGGGGCTGC
 AGCGCTACCTCCAGGGGAGCTCCGTGCAGCTCAGGAGCCTCCGTCTCACCTATCGCAACCTATCG
 GGCCCTGATAAGCGGCTGGTGACGCTGCGACTGCTGCTCGCTCGCTGAGTACACGGTCACCCA
 GCTGCGGCCCAACGCCACTTACTCCGTCTGTGTCATGCCCTTTGGGGCCCGGGCGGGTGCCGGAGG
 GCGAGGAGCCTGCGGGGAGGCCATACACCCCGAGCCGTCCACTCCAACACGCCCCAGTCACC
 CAGGCCCGCGAGGGCAACCTGCCGCTCCTCATTGCGCCCGCCCTGGCCGCGGTGCTCCTGGCCGC
 GCTGGCTGCGGTGGGGGACGCTACTGTGTGCGGCGGGGGCGGGCCATGGCAGCAGCGGCTCAGG
 ACAAGGGCAGGTGGGGCCAGGGGCTGGGCCCTTGAACTGGAGGGAGTGAAGGTCCCTTGAGAG
 CCAGGCCCGAAGGCAACAGAGGGCGGTGGAGAGGCCCTGCCAGCGGGTCTGAGTGTGAGGTGCC
 ACTCATGGGCTTCCAGGGCCTGGCCTCCAGTCACCCCTCCACGCAAAGCCCTACATCTAAGCCA
 GAGAGAGACAGGGCAGCTGGGGCCGGGCTCTCAGCCAGTGAGATGGCCAGCCCCCTCCTGCTGCC
 ACACCACGTAAGTTCTCAGTCCCAACCTCGGGGATGTGTGCAGACAGGGCTGTGTGACCACAGCT
 GGGCCCTGTTCCCTCTGGACCTCGGTCTCCTCATCTGTGAGATGCTGTGGCCAGCTGACGAGCC
 CTAACGTCCCCAGAACCAGTGCCATGAGGACAGTGTCCGCCCTGCCCTCCGCAACGTGCAGTC
 CCTGGGCACGGCGGGCCCTGCCATGTGCTGGTAACGCATGCCTGGGTCTGCTGGGCTCTCCAC
 TCCAGGCGGACCCCTGGGGCCAGTGAAGGAAGCTCCCGGAAAGAGCAGAGGGAGAGCGGGTAGGC
 GGCTGTGTGACTCTAGTCTTGGCCCCAGGAAGCGAAGGAACAAAAGAACTGGAAAGGAAGATGC
 TTTAGGAACATGTTTTGCTTTTTTAAATATATATATTTATAAGAGATCCTTTCCCATTTATTCT
 GGGAAAGATGTTTTCAAACCTCAGAGACAAGGACTTTGGTTTTTGTAAAGACAAACGATGATATGAA
 GGCCTTTTGTAAAGAAAAATAAAGATGAAGTGTGAAA

FIGURE 16

MCSRVP LLLPL LLL L L L L L L A L G P G V Q G C P S G C Q C S Q P Q T V F C T A R Q G T T V P R D V P P D T V G L Y V F E N G I T
M L D A G S F A G L P G L Q L L D L S Q N Q I A S L P S G V F Q P L A N L S N L D L T A N R L H E I T N E T F R G L R R L E R L Y
L G K N R I R H I Q P G A F D T L D R L L E L K L Q D N E L R A L P P L R L P R L L L L D L S H N S L L A L E P G I L D T A N V E
A L R L A G L G L Q Q L D E G L F S R L R N L H D L D V S D N Q L E R V P P V I R G L R G L T R L R L A G N T R I A Q L R P E D L
A G L A A L Q E L D V S N L S L Q A L P G D L S G L F P R L R L L A A A R N P F N C V C P L S W F G P W V R E S H V T L A S P E E
T R C H F P P K N A G R L L L E L D Y A D F G C P A T T T T A T V P T T R P V V R E P T A L S S S L A P T W L S P T A P A T E A P
S P P S T A P P T V G P V P Q P Q D C P P S T C L N G G T C H L G T R H H L A C L C P E G F T G L Y C E S Q M G Q G T R P S P T P
V T P R P P R S L T L G I E P V S P T S L R V G L Q R Y L Q G S S V Q L R S L R L T Y R N L S G P D K R L V T L R L P A S L A E Y
T V T Q L R P N A T Y S V C V M P L G P R V P E G E E A C G E A H T P P A V H S N H A P V T Q A R E G N L P L L I A P A L A A V
L L A A L A A V G A A Y C V R R G R A M A A A A Q D K G Q V G P G A G P L E L E G V K V P L E P G P K A T E G G G E A L P S G S E
C E V P L M G F P G P G L Q S P L H A K P Y I

Important features:

Signal peptide:

amino acids 1-23

Transmembrane domain:

amino acids 579-599

EGF-like domain cysteine pattern signature.

amino acids 430-442

Leucine zipper pattern.

amino acids 197-219, 269-291

N-glycosylation sites.

amino acids 101-105, 117-121, 273-277, 500-504, 528-532

Tyrosine kinase phosphorylation sites.

amino acids 124-131, 337-345

N-myristoylation sites.

amino acids 23-29, 27-33, 70-76, 142-148, 187-193, 348-354,
594-600, 640-646

FIGURE 17

GCAGCGGCGAGGCGGGCGGTGGTGGCTGAGTCCGTGGTGGCAGAGGCGAAGGCGACAGCTCATGCG
GGTCCGGATAGGGCTGACGCTGCTGCTGTGTGCGGTGCTGCTGAGCTTGGCCTCGGCGTCTCGG
ATGAAGAAGGCAGCCAGGATGAATCCTTAGATTCCAAGACTACTTTGACATCAGATGAGTCAGTA
AAGGACCATACTACTGCAGGCAGAGTAGTTGCTGGTCAAATATTTCTTGATTGAGAAGAATCTGA
ATTAGAATCCTCTATTCAAGAAGAGGAAGACAGCCTCAAGAGCCAAGAGGGGAAAGTGTACAG
AAGATATCAGCTTTCTAGAGTCTCCAAATCCAGAAAACAAGGACTATGAAGAGCCAAAGAAAGTA
CGGAAACCAGCTTTGACCGCCATTGAAGGCACAGCACATGGGGAGCCCTGCCACTTCCCTTTTCT
TTTCCTAGATAAGGAGTATGATGAATGTACATCAGATGGGAGGGAAGATGGCAGACTGTGGTGTG
CTACAACCTATGACTACAAAGCAGATGAAAAGTGGGGCTTTTGTGAACTGAAGAAGAGGCTGCT
AAGAGACGGCAGATGCAGGAAGCAGAAATGATGTATCAAATGGAATGAAAATCCTTAATGGAAG
CAATAAGAAAAGCCAAAAAGAGAAGCATATCGGTATCTCCAAAAGGCAGCAAGCATGAACCATA
CCAAAGCCCTGGAGAGAGTGCATATGCTCTTTTATTTGGTGATTACTTGCCACAGAATATCCAG
GCAGCGAGAGAGATGTTTGAGAAGCTGACTGAGGAAGGCTCTCCCAAGGGACAGACTGCTCTTGG
CTTTCTGTATGCCTCTGGACTTGGTGTAAATCAAGTCAGGCAAAGGCTCTTGATATTATACAT
TTGGAGCTCTTGGGGGAATCTAATAGCCACATGGTTTTGGTAAGTAGACTTTAGTGGAAGGCT
AATAATATTAACATCAGAAGAATTTGTGGTTTATAGCGGCCACAACTTTTTCAGCTTTCATGATC
CAGATTTGCTTGATTATTAAGACCAAATATTCAGTTGAACCTCCTTCAAATTCTTGTTAATGGATAT
AACACATGGAATCTACATGTAAATGAAAGTTGGTGGAGTCCACAATTTTTCTTTAAATGATTAG
TTTGCTGATTGCCCTAAAAAGAGAGATCTGATAAATGGCTCTTTTAAATTTTCTCTGAGTTG
GAATTGTCAGAATCATTTTTTACATTAGATTATCATAATTTTAAAAATTTTCTTTAGTTTTTCA
AAATTTGTAAATGGTGGCTATAGAAAACAACATGAAATATTATACAATATTTTGCAACAATGC
CCTAAGAATTGTTAAATTCATGGAGTTATTTGTGCAGAAATGACTCCAGAGAGCTCTACTTTCTG
TTTTTTACTTTTCATGATTGGCTGTCTTCCCATTTATTCTGGTCATTTATTGCTAGTGACACTGT
GCCTGCTTCCAGTAGTCTCATTTTCCCTATTTTGCTAATTTGTTACTTTTTCTTTGCTAATTTGG
AAGATTAACCTCATTTTAAATAAAATATGTCTAAGATTAAAAA
AAA

FIGURE 18

MRVRIGLTLLLCVLLSLASASSDEEGSQDESLSKTTLTSDSVKDHTTAGRVVAGQIFLDSESESEL
ESSIQEEEDSLKSQEGESVTEDISFLESPNPENKDYEEPKKVRKPALTAIEGTAHGEPCHFPFLFLDK
EYDECTSDGREDGRLWCATTYDYKADEKWFCEEEEEAAKRRQMQEAEEMYQTGMKILNGSNKKSQKR
EAYRYLQKAASMNHTKALERVSYALLFGDYLPQNIQAAREMFEEKLTEEGSPKGQTALGFLYASGLGVN
SSQAKALVYYTFGALGGNLIAMVLVSRL

Important features:

Signal peptide:

amino acids 1-21

N-glycosylation sites.

amino acids 195-199, 217-221, 272-276

Tyrosine kinase phosphorylation site.

amino acids 220-228

N-myristoylation sites.

amino acids 120-126, 253-259, 268-274, 270-274, 285-291, 289-295

Glycosaminoglycan attachment site.

amino acids 267-271

Microbodies C-terminal targeting signal.

amino acids 299-303

Type II fibronectin collagen-binding domain protein.

amino acids 127-169

Fructose-bisphosphate aldolase class-II protein.

amino acids 101-119

FIGURE 19

AATTCAGATTTTAAGCCCATCTGCAGTGAATTTTCATGAAC TAGCAAGAGGACACCATCTTCTT
GTATTATACAAGAAAGGAGTGACCTATCACACACAGGGGAAAAATGCTCTTTTGGGTGCTAGG
CCTCCTAATCCTCTGTGGTTTTCTGTGGACTCGTAAAGGAAAATAAGATTGAAGACATCACTG
ATAAGTACATTTTTATCACTGGATGTGACTCGGGCTTTGGAACTTGGCAGCCAGAACCTTTGAT
AAAAAGGGATTTCATGTAATCGCTGCCTGTCTGACTGAATCAGGATCAACAGCTTTAAAGGCAGA
AACCTCAGAGAGACTTCGTACTGTGCTTCTGGATGTGACCGACCCAGAGAATGTCAAGAGGACTG
CCCAGTGGGTGAAGAACCAAGTTGGGGAGAAAGTCTCTGGGGTCTGATCAATAATGCTGGTGTT
CCCGGCGTGCTGGCTCCCACTGACTGGCTGACACTAGAGGACTACAGAGAACCATTGAAGTGAA
CCTGTTTGGACTCATCAGTGTGACACTAAATATGCTTCCTTTGGTCAAGAAAGCTCAAGGGAGAG
TTATTAATGTCTCCAGTGTGGAGGTCGCCTTGCAATCGTTGGAGGGGGCTATACTCCATCCAAA
TATGCAGTGAAGGTTTCAATGACAGCTTAAGACGGGACATGAAAGCTTTTGGTGTGCACGCTC
ATGCATTGAACCAGGATTGTTCAAAACAACTTGGCAGATCCAGTAAAGGTAATTGAAAAAAAC
TCGCCATTTGGGAGCAGCTGTCTCCAGACATCAAACAACAATATGGAGAAGGTTACATTGAAAA
AGTCTAGACAACTGAAAGGCAATAAATCCTATGTGAACATGGACCTCTCTCCGGTGGTAGAGTG
CATGGACCACGCTCTAACAAGTCTCTTCCCTAAGACTCATTATGCCGCTGGAAAAGATGCCAAA
TTTTCTGGATACCTCTGTCTCACATGCCAGCAGCTTTGCAAGACTTTTTATTGTTGAAACAGAAA
GCAGAGCTGGCTAATCCCAAGGCAGTGTGACTCAGCTAACCACAAATGTCTCCTCCAGGCTATGA
AATTGGCCGATTTCAAGAACACATCTCCTTTTCAACCCCATTCCTTATCTGCTCCAACCTGGACT
CATTTAGATCGTGCTTATTTGGATTGCAAAAGGGAGTCCCACCATCGCTGGTGGTATCCCAGGGT
CCCTGCTCAAGTTTTCTTTGAAAAGGAGGGCTGGAATGGTACATCACATAGGCAAGTCTGCCCT
GTATTTAGGCTTTGCCTGCTTGGTGTGATGTAAGGGAAATTGAAAGACTTGCCCATTCAAAATGA
TCTTTACCGTGGCCTGCCCCATGCTTATGGTCCCAGCATTTACAGTAACTTGTGAATGTTAAGT
ATCATCTCTTATCTAAATATTTAAAGATAAGTCAACCCAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAA

FIGURE 20

MLFWVLGLLILCGFLWTRKGLKIEDITDKYIFITGCDSGFGNLAARTFDKKGFHVIAACLTESG
STALKAETSERLRTVLLDVTDPENVKRTAQWVKNOVGEKGLWGLINNAGVPGVLAPTDWLTLEDY
REPIEVNLFGLISVTLNMLPLVKKAQGRVINVSSVGGRLAIVGGGYTPSKYAVEGFNDLRRDMK
AFGVHVSCEPGLFKTNLADPVKVIEKKLAIWEQLSPDIKQQYGEgyIEKSLDKLKGKNSYVNMD
LSPVVECMDHALTSLFPKTHYAAGKDAKIFWIPLSHMPAALQDFLLLKQKAELANPKAV

Important features of the protein:

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 136-152

N-glycosylation sites.

amino acids 161-163, 187-190 and 253-256

Glycosaminoglycan attachment site.

amino acids 39-42

N-myristoylation sites.

amino acids 36-41, 42-47, 108-113, 166-171, 198-203 and 207-212

FIGURE 21

CTGAGGCGGCGGTAGCATGAGGGGGAGAGTACGTGCGGCGGTGCTCTCGGGCTTTGTGCTCGGCG
CACTCGCTTTCCAGCACCTCAACACGGACTCGGACACGGAAGGTTTTCTTCTGGGGAAGTAAAA
GGTGAAGCCAAGAACAGCATTACTGATTCCCAAATGGATGATGTTGAAGTTGTTTATACAATTGA
CATTCAGAAATATATTCCATGCTATCAGCTTTTTAGCTTTTATAATTCTTCAGGCGAAGTAAATG
AGCAAGCACTGAAGAAAATATTATCAATGTCAAAAAGAATGTGGTAGGTTGGTACAAATCCGT
CGTCATTAGATCAGATCATGACGTTTAGAGAGAGGCTGCTTCACAAAACTTGCAGGAGCATTT
TTCAAACCAAGACCTTGTTTTCTGCTATTAAACCAAGTATAATAACAGAAAGCTGCTCTACTC
ATCGACTGGAACATTCTTTATATAAACCTCAAAAAGGACTTTTTTCACAGGGTACCTTTAGTGGTT
GCCAATCTGGGCATGTCTGAACAAGTGGTTATAAACTGTATCAGGTTCTGTATGTCCACTGG
TTTTAGCCGAGCAGTACAAACACACAGCTCTAAATTTTTGAAGAAGATGGATCCTTAAAGGAGG
TACATAAGATAAATGAAATGTATGCTTCATTACAAGAGGAATTAAAGAGTATATGCAAAAAAGTG
GAAGACAGTGAACAAGCAGTAGATAAACTAGTAAAGGATGTAAACAGATTAAACGAGAAATTGA
GAAAAGGAGAGGAGCACAGATTCAGGCAGCAAGAGAGAAGAACATCCAAAAGACCCCTCAGGAGA
ACATTTTTCTTTGTCAGGCATTACGGACCTTTTTTCCAAATCTGAATTTCTTCATTCATGTGTT
ATGTCTTTAAAAATAGACATGTTTCTAAAAGTAGCTGTAACACACCACCATCTCGATGTAGT
AGACAATCTGACCTTAATGGTAGAACACACTGACATTCCTGAAGCTAGTCCAGCTAGTACACCAC
AAATCATTAAGCATAAAGCCTTAGACTTAGATGACAGATGGCAATTCAGAGATCTCGGTTGTTA
GATACACAAGACAAACGATCTAAAGCAAATACTGGTAGTAGTAACCAAGATAAAGCATCCAAAT
GAGCAGCCCAGAAACAGATGAAGAAATTGAAAAGATGAAGGGTTTTGGTGAATATTCACGGTCTC
CTACATTTGATCCTTTTAACCTTACAAGGAGATTTTTTTATTTGGCTGATGGGTAAAGCCAAAC
ATTTCTATTGTTTTTACTATGTTGAGCTACTTGCAGTAAGTTCATTTGTTTTTACTATGTTACAC
TGTTTGCGAGTAATACACAGATAACTCTTAGTGCATTTACTTCACAAAGTACTTTTTCAAACATCA
GATGCTTTTATTTCCAAACCTTTTTTTCACCTTTCACCTAAGTTGTTGAGGGGAAGGCTTACACAG
ACACATCTTTAGAATTGGAAAAGTGAGACCAGGCACAGTGGCTCACACCTGTAATCCCAGCACT
TAGGGAAGACAAGTCAGGAGGATTGATTGAAGCTAGGAGTTAGAGACCAGCCTGGGCAACGTATT
GAGACCATGTCTATTAAAAAATAAAATGGAAAAGCAAGAATAGCCTTATTTTCAAATATGGAAA
GAAATTTATATGAAATTTATCTGAGTCATTAAATTTCTCCTTAAGTGATACTTTTTTAGAAGTA
CATTATGGCTAGAGTTGCCAGATAAAATGCTGGATATCATGCAATAAATTTGCAAAACATCATCT
AAAATTTAAAAAAAAAAAAAAAAAAAAA

FIGURE 22

MEGESTSAVLSGFVLGALAFQHLNTDSDTEGFLLGEVKGEAKNSITDSQMDDVEVVYTTIDIQKYI
PCYQLFSFYNSSGEVNEQALKKILSNVKKNVVGWYKFRRHSDQIMTFRERLLHKNLQEHFSNQDL
VFLLLTPSIITESCSTHRLEHSLYKPQKGLFHRVPLVVANLGMSEQLGYKTVSGSCMSTGFSRAV
QTHSSKFFEEEDGSLKEVHKINEMYASLQEELKSICKKVEDSEQAVDKLVKDVNRLKREIEKRRGA
QIQAAREKNIQKDPQENIFLCQALRTFFPNSEFLHSCVMSLKNRHVSKSSCNYNHHLDVVDNLTL
MVEHTDIPEASPASTPQIIKHKALDLDLRWQFKRSRLDQDKRSKANTGSSNQDKASKMSSPET
DEEIEKMKGFGEYSRSPTF

Important features:

Signal peptide:

amino acids 1-19

N-glycosylation sites.

amino acids 75-79, 322-326

N-myristoylation site.

amino acids 184-154

Growth factor and cytokines receptors family.

amino acids 134-150

FIGURE 23

GGCACAGCCGCGCGGCGAGGGCAGAGTCAGCCGAGCCGAGTCCAGCCGGACGAGCGGACCAGCGAGGGCAGCCCAA
GCAGCGCGCAGCGAACGCCGCCGCCGCCACACCTCTGCGGTCCCCGCGGCGCCTGCCACCTTCCCTCCTTCCCC
GCGTCCCCGCTCGCCGCCAGTCAGCTTGCCTGGGTTGCTGCCCCGCGAAACCCGAGGTACACAGCCCGCGCCTCT
GCTTCCCTGGGCCGCGCGCCCTCCACGCCCTCTTCTCCCTGGCCGCGCCTGGCACCGGGGACCGTTGCCTGA
CGCGAGGCCCAGCTCTACTTTTCGCCCGCGCTCCTCCGCTGCTCGCCTCTTCCACCACTCCAACCTCTTCTCC
TCCAGCTCCACTCGTAGTCCCCGACTCCGCCAGCCCTCGGCCGCTGCGTAGCGCGCTTCCCGTCCGGTCCCAAA
GGTGGGAACGCGTCCGCCCGGCCCGCACCATGGCACGGTTGCGCTTGGCCGCGCTTCTGTCACCTGGCAGTGCTC
AGCGCGCGCTGCTGGTGGCAGCTCAAGTCGAAAAGTTGCTCGGAAGTGCAGCTCTTACGTGTCAAAGGCTTC
AACAGAAGCATGCCCCCTCCACGAGATCAACGGTGATCATTTGAAGATCTGTCCCCAGGGTTCTACCTGCTGCTCT
CAAGAGATGGAGGAGAAGTACAGCTGCAAAAGTAAAGATGATTTCAAAAGTGTGGTCAGCGAACAGTGCAATCATTTG
CAAGCTGTCTTTGCTTACGTTACAAGAGTTTGTATGAATCTTCAAGAAGTACTTGAAGTGCAGAGAAATCCCTG
AATGATATGTTGTGAAGACATATGGCCATTTATACATGCAAAATCTGAGCTATTTAAAGATCTCTCGTAGAGTTG
AAACGTTACTACGTGGTGGGAAATGTGAACCTGGAAGAAATGCTAAATGACTTCTGGGCTCGCCTCCTGGAGCGGATG
TTCCGCTGCTGAAGTCCAGTACCACTTTACAGATGAGTATCTGGAATGTGTGAGCAAGTATACGAGCAGCTGAAG
CCCTTCGGAGATGTCCTTCGCAAAATTGAAGCTCCAGGTTACTCGTGCTTTTGTAGCAGCCCGTACTTTGCTCAAGGC
TTAGCGGTTGCGGGAGATGTCGTGAGCAAGGTCTCCGTGGTAAACCCACAGCCAGTGTACCCATGCCCTGTTGAAG
ATGATCTACTGCTCCCACTGCCGGGTCTCGTGACTGTGAAGCCATGTTACAAGTACTGCTCAAACATCATGAGAGGC
TGTTTGGCCAACCAAGGGGATCTCGATTTTGAATGGAACAATTTATAGATGCTATGCTGATGGTGGCAGAGAGGCTA
GAGGGTCTTTCAACATTGAATCGGTGATGATCCCATCGATGTGAAGATTTCTGATGCTATTATGAACATGCAGGAT
AATAGTGTTCAGTGTCTCAGAAGGTTTTCAGGGATGTGGACCCCAAGCCCTCCAGCTGGACGAATTTCTCGT
TCCATCTCTGAAAGTGCCCTTCAGTGCTCGCTTCAGACCACATCACCCGAGGAACGCCAACACAGCAGCTGGCACT
AGTTTGGACCGACTGGTTACTGATGTCAAGGAGAACTGAAACAGGCCAAGAAATCTGGTCTCCCTTCGAGCAAC
GTTTGAACGATGAGAGGATGGCTGCAGGAAACGGCAATGAGGATGACTGTTGGAATGGGAAAGGCAAAAGCAGGTAC
CTGTTTGCAGTGACAGGAAATGGATTAGCCAACAGGGCAACAACCCAGAGGTCCAGGTTGACACCAGCAAACAGAC
ATACTGATCCTTCGTCAAATCATGGCTCTTCGAGTGATGACCAGCAAGATGAAGAATGCATACAATGGGAACGACGTG
GACTTCTTTGATATCAGTGATGAAAGTAGTGGAGAAGGAAGTGGAAAGTGGCTGTGAGTATCAGCAGTGCCCTTCAGAG
TTTGACTACAATGCCACTGACCATGTGGGAAGAGTGCCAATGAGAAAGCCGACAGTGCTGGTGTCCGTCTGGGGCA
CAGGCCCTACCTCCTCACTGTCTTCTGCATCTTGTTCCTGGTTATGCAGAGAGAGTGGAGATAATTCTCAAACCTCTGAG
AAAAAGTGTTTCATCAAAAAGTTAAAAGGCACAGTTATCACTTTTCTACCATCCTAGTGACTTTGCTTTTAAATGAA
TGGACAACAATGTACAGTTTTTACTATGTGGCACTGGTTTAAAGAGTGTGACTTTGTTTTCTCATTCACTTTTGGG
AGGAAAAGGGACTGTGCATTGAGTTGGTTCTGCTCCCCAAACCATGTTAAACGTGGCTAACAGGTAGGTACAGAA
CTATAGTTAGTTGTGATTTGTGATTTTATCACTCTATTTTGTGTTGATGTTTTTCTCATTTCGTTTGTGGGT
TTTTTTTCCAACGTGTGATCTCGCTTGTTCCTTACAAGCAAACAGGGTCCCTTCTTGGCACGTAACATGTACGTATT
TCTGAAATATTAAATAGCTGTACAGAAGCAGGTTTTTATTTATCATGTTATCTTATTAAGAAAAGCCCAAAAGC

FIGURE 24

MARFGLPALLCTLAVLSAALLAAELKSKSCSEVRRLYVSKGFNKNDAPLHEINGDHLKICPQGST
CCSQEMEKEYSLQSKDDFKSVVSEQCNHLQAVFASRYKKFDEFFKELLENAEKSLNDMFVKTYGH
LYMQNSELFKDLFVELKRYVVGNVNLEEMLNDFWARLLERMFRVLVNSQYHFTDEYLECVSKYTE
QLKPFQDVPRKLKLQVTRAFVAARTFAQGLAVAGDVVSKVSVVNPTAQCTHALLKMIYCSHCRGL
VTVKPCYNYCSNIMRGCLANQGDLD FEWNNFIDAMLMVAERLEGPFNIESVMDPIDVKISDAIMN
MQDNSVQVSQKVFQCGPPKPLPAGRISRSISESAFSARFRPHHPEERPTTAAGTSLDRLVTDVK
EKLKQAKKFWSSLP SNVCNDEMAAGNGNEDDCWNGKGKSRYLFAVTGNGLANQGNNEVQVDT
KPDILILRQIMALRVMTSKMKNAYNNGNDVDFDISDESSGEGSGSGCEYQQCPSEFDYNATDHAG
KSANEKADSAGVRPGAQAYLLTVFCILFLVMQREWR

Important features:

Signal peptide:

amino acids 1-22

ATP/GTP-binding site motif A (P-loop).

amino acids 515-524

N-glycosylation site.

amino acids 514-518

Glycosaminoglycan attachment sites.

amino acids 494-498, 498-502

N-myristoylation sites.

amino acids 63-69, 224-230, 276-282, 438-444, 497-503, 531-537

Glypicans proteins.

amino acids 54-75, 105-157, 238-280, 309-346, 423-460, 468-506

FIGURE 25

CTCGCCCTCAAATGGGAACGCTGGCCTGGGACTAAAGCATAGACCACCAGGCTGAGTATCCTGAC
CTGAGTCATCCCCAGGGATCAGGAGCCTCCAGCAGGGAACCTTCCATTATATTCTTCAAGCAACT
TACAGCTGCACCGACAGTTGCGAATGAAAGTTCTAATCTCTTCCCTCCTCCTGTTGCTGCCACTAA
TGCTGATGTCCATGGTCTCTAGCAGCCTGAATCCAGGGGTCGCCAGAGGCCACAGGGACCGAGGC
CAGGCTTCTAGGAGATGGCTCCAGGAAGGCGGCCAAGAATGTGAGTGCAAAGATTGGTTCCCTGAG
AGCCCCGAGAAGAAAATTTCATGACAGTGTCTGGGCTGCCAAAGAAGCAGTGCCCTGTGATCATT
TCAAGGGCAATGTGAAGAAAACAAGACACCAAAGGCACCACAGAAAGCCAAACAAGCATTCCAGA
GCCTGCCAGCAATTTCTCAAACAATGTCAGCTAAGAAGCTTTGCTCTGCCTTTGTAGGAGCTCTG
AGCGCCCACTCTTCCAATTAAACATTCTCAGCCAAGAAGACAGTGAGCACACCTACCAGACACTC
TTCTTCTCCACCTCACTCTCCCACTGTACCCACCCCTAAATCATTCCAGTGCTCTCAAAAAGCA
TGTTTTCAAGATCATTTTGTGTTGCTCTCTCTAGTGTCTTCTCTCTCGTCAGTCTTAGCCT
GTGCCCTCCCCTTACCCAGGCTTAGGCTTAATTACCTGAAAGATTCCAGGAAACTGTAGCTTCCT
AGCTAGTGTCAATTAACCTTAAATGCAATCAGGAAAGTAGCAAACAGAAGTCAATAAATATTTTT
AAATGTCAAAAAAAAAAAAAAAAAA

FIGURE 26

MKVLISSLLLLLPLMLMSMVSSSLNPGVARGHRDRGQASRRWLQEGGQECECKDWFLRAPRRKFM
TVSGLPKKQCPCDHFKGNVKKTRHQRHHRKPNKHSRACQQFLKQCQLRSFALPL

Important features:

Signal peptide:

amino acids 1-22

N-myristoylation sites.

amino acids 27-33, 46-52

FIGURE 27

GGACGCCAGCGCCTGCAGAGGCTGAGCAGGGAAAAAGCCAGTGCCCCAGCGGAAGCACAGCTCAG
AGCTGGTCTGCCATGGACATCCTGGTCCCCTCCTGCAGCTGCTGGTGCTGCTTCTTACCCTGCC
CCTGCACCTCATGGCTCTGCTGGGCTGCTGGCAGCCCCTGTGCAAAAGCTACTTCCCCTACCTGA
TGCCCGTGCTGACTCCCAAGAGCAACCGCAAGATGGAGAGCAAGAAACGGGAGCTCTTCAGCCAG
ATAAAGGGGCTTACAGGAGCCTCCGGGAAAGTGGCCCTACTGGAGCTGGGCTGCGGAACCGGAGC
CACTTTTCAGTTCTACCCACCGGGCTGCAGGGTCACCTGCCTAGACCCAAATCCCCACTTTGAGA
AGTTCCTGACAAAGAGCATGGCTGAGAACAGGCACCTCCAATATGAGCGGTTTGTGGTGGCTCCT
GGAGAGGACATGAGACAGCTGGCTGATGGCTCCATGGATGTGGTGGTCTGCACTCTGGTGCTGTG
CTCTGTGCAGAGCCCAAGGAAGGTCTGCAGGAGGTCCGGAGAGTACTGAGACCGGGAGGTGTGC
TCTTTTCTGGGAGCATGTGGCAGAACCATATGGAAGCTGGGCCTTCATGTGGCAGCAAGTTTTC
GAGCCACCTGGAAACACATTGGGGATGGCTGCTGCCTCACCAGAGAGACCTGGAAGGATCTTGA
GAACGCCAGTTCTCCGAAATCCAATGGAACGACAGCCCCCTCCCTTGAAGTGGCTACCTGTTG
GGCCCCACATCATGGGAAAGGCTGTCAAACAATCTTTCCCAAGCTCCAAGGCACTCATTTGCTCC
TTCCCCAGCCTCCAATTAGAACAAGCCACCCACCAGCCTATCTATCTTCCACTGAGAGGGACCTTA
GCAGAATGAGAGAAGACATTTCATGTACCACCTACTAGTCCCTCTCTCCCCAACCTCTGCCAGGGC
AATCTCTAACTTCAATCCCGCCTTCGACAGTGAAAAGCTCTACTTCTACGCTGACCCAGGGAGG
AAACACTAGGACCCTGTTGTATCCTCAACTGCAAGTTTCTGGACTAGTCTCCCAACGTTTGCCTC
CCAATGTTGTCCCTTTCTTCCGTTCCCATGGTAAAGCTCCTCTCGCTTTCTCCTGAGGCTACAC
CCATGCGTCTCTAGGAAGTGGTCACAAAAGTCATGGTGCCTGCATCCCTGCCAAGCCCCCTGAC
CCTCTCTCCCCACTACCACCTTCTTCTGAGCTGGGGGACCAGGGAGAATCAGAGATGCTGGGG
ATGCCAGAGCAAGACTCAAAGAGGCAGAGGTTTGTCTCAAATATTTTAAATAAATAGACGAA
ACCACG

FIGURE 28

MDILVPLLQLLVLLLTPLPLHLMALLGCWQPLCKSYFPYLMAVLTPKSNRKMESKKRELFSSQIKGL
TGASGKVALLELGCGTGANFQFYPPGCRVTCIDPNPHFEKFLTKSMAENRHLQYERFVVAPGEDM
RQLADGSMDVVVCTLVLCVQSPRKVLQEVRRVLRPGGVLEFFWEHVAEPYGSWAFMWQQVFEPWTW
KHIGDGCCLTRETWKDLENAQFSEIQMERQPPPLKWLVPVGPVHMGKAVKQSFSSKALICSFPSL
QLEQATHQPIYLPRLGT

Important features:

Signal peptide:

amino acids 1-23

Leucine zipper pattern.

amino acids 10-32

N-myristoylation sites.

amino acids 64-70, 78-84, 80-86, 91-97, 201-207

FIGURE 29

CAATGTTTGCCTATCCACCTCCCCAAGCCCCTTTACCTATGCTGCTGCTAACGCTGCTGCTGCT
GCTGCTGCTGCTTAAAGGCTCATGCTTGGAGTGGGGACTGGTCGGTGCCAGAAAGTCTCTTCTG
CCACTGACGCCCCCATCAGGGATTGGGCCTTCTTTCCCCCTTCCTTTCTGTGTCTCCTGCCTCAT
CGGCCTGCCATGACCTGCAGCCAAGCCCAGCCCCGTGGGGAAGGGGAGAAAGTGGGGGATGGCTTA
AGAAAGCTGGGAGATAGGGAACAGAAGAGGGTAGTGGGTGGGCTAGGGGGGCTGCCTTATTTAAA
GTGGTTGTTTATGATTCTTATACTAATTTATACAAAGATATTAAGGCCCTGTTTATTAAGAAATT
GTTCCCTTCCCCCTGTGTTCAATGTTTGTAAAGATTGTTCTGTGTAAATATGTCTTTATAATAAAC
AGTTAAAAGCTGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

1970	1971	1972	1973	1974	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100	2101	2102	2103	2104	2105	2106	2107	2108	2109	2110	2111	2112	2113	2114	2115	2116	2117	2118	2119	2120	2121	2122	2123	2124	2125	2126	2127	2128	2129	2130	2131	2132	2133	2134	2135	2136	2137	2138	2139	2140	2141	2142	2143	2144	2145	2146	2147	2148	2149	2150	2151	2152	2153	2154	2155	2156	2157	2158	2159	2160	2161	2162	2163	2164	2165	2166	2167	2168	2169	2170	2171	2172	2173	2174	2175	2176	2177	2178	2179	2180	2181	2182	2183	2184	2185	2186	2187	2188	2189	2190	2191	2192	2193	2194	2195	2196	2197	2198	2199	2200	2201	2202	2203	2204	2205	2206	2207	2208	2209	2210	2211	2212	2213	2214	2215	2216	2217	2218	2219	2220	2221	2222	2223	2224	2225	2226	2227	2228	2229	2230	2231	2232	2233	2234	2235	2236	2237	2238	2239	2240	2241	2242	2243	2244	2245	2246	2247	2248	2249	2250	2251	2252	2253	2254	2255	2256	2257	2258	2259	2260	2261	2262	2263	2264	2265	2266	2267	2268	2269	2270	2271	2272	2273	2274	2275	2276	2277	2278	2279	2280	2281	2282	2283	2284	2285	2286	2287	2288	2289	2290	2291	2292	2293	2294	2295	2296	2297	2298	2299	2300	2301	2302	2303	2304	2305	2306	2307	2308	2309	2310	2311	2312	2313	2314	2315	2316	2317	2318	2319	2320	2321	2322	2323	2324	2325	2326	2327	2328	2329	2330	2331	2332	2333	2334	2335	2336	2337	2338	2339	2340	2341	2342	2343	2344	2345	2346	2347	2348	2349	2350	2351	2352	2353	2354	2355	2356	2357	2358	2359	2360	2361	2362	2363	2364	2365	2366	2367	2368	2369	2370	2371	2372	2373	2374	2375	2376	2377	2378	2379	2380	2381	2382	2383	2384	2385	2386	2387	2388	2389	2390	2391	2392	2393	2394	2395	2396	2397	2398	2399	2400	2401	2402	2403	2404	2405	2406	2407	2408	2409	2410	2411	2412	2413	2414	2415	2416	2417	2418	2419	2420	2421	2422	2423</
------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	--------

Important features:

Signal peptide:

Growth factor and cytokines receptors family:

amino acids 3-18

FIGURE 31

GTTTGAATTCCTTCAACTATACCCACAGTCCAAAAGCAGACTCACTGTGTCCCAGGCTACCAAGTT
CCTCCAAGCAAGTCATTTCCCTTATTTAACCGATGTGTCCCTCAAACACCTGAGTGCTACTCCCT
ATTTGCATCTGTTTTGATAAATGATGTTGACACCCCTCCACCGAATTCCTAAGTGGAATCATGTCGG
GAAGAGATACAATCCTTGGCCTGTGTATCCTCGCATTAGCCTTGTCTTTGGCCATGATGTTTACC
TTCAGATTCAACACCCCTTCTGGTTCACATTTTCATTTATTGTTTATTTTGGGATTGTTGTT
TGTCTGCGGTGTTTTATGGTGGCTGTATTATGACTATAACACGACCTCAGCATAGAATTGGACA
CAGAAAGGGGAAAATATGAAGTGCCTGCTGGGGTTTGCTATCGTATCCACAGGCATCACGGCAGTG
CTGCTCGTCTTGATTTTTGTTCTCAGAAAGAGAATAAAATTGACAGTTGAGCTTTTCCAAATCAC
AAATAAAGCCATCAGCAGTGCTCCCTTCCTGCTGTTCCAGCCACTGTGGACATTTGCCATCCTCA
TTTTCTTCTGGGTCTCTGGGTGGCTGTGCTGCTGAGCCTGGGAAGTGCAGGAGCTGCCAGGTT
ATGGAAGGCGGCCAAGTGGAATATAAGCCCCCTTCGGGCATTCCGGTACATGTGGTCGTACCATTT
AATTGGCCTCATCTGGACTAGTGAATTCATCCTTGCGTGCCAGCAAATGACTATAGCTGGGGCAG
TGGTTACTTGTATTTTCAACAGAAGTAAAAATGATCCTCCTGATCATCCCATCCTTTTCGTCTCTC
TCCATTCTCTTCTTCTACCATCAAGGAACCGTTGTGAAAGGGTCATTTTAAATCTCTGTGGTGAG
GATTCCGAGAATCATTGTGATGTACATGAAAACGCACTGAAAGAACAGCAGCATGGTGCATTGT
CCAGGTACCTGTTCCGATGCTGCTACTGCTGTTTCTGGTGTCTTGACAAATACCTGCTCCATCTC
AACCAGAATGCATATACTACAACCTGCTATTAATGGGACAGATTTCTGTACATCAGCAAAAGATGC
ATTCAAAATCTTGTCCAAGAACTCAAGTCACCTTACATCTATTAAGTCTTTGGAGACTTCATAA
TTTTTCTAGGAAAGGTGTTAGTGGTGTGTTTCACTGTTTTTGGAGGACTCATGGCTTTTAACTAC
AATCGGGCATTCCAGGTGTGGGCAGTCCCTCTGTTATTGGTAGCTTTTTTGCCTACTTAGTAGC
CCATAGTTTTTTATCTGTGTTTGAAACTGTGCTGGATGCACTTTTCTGTGTTTTGCTGTTGATC
TGGAACAAATGATGGATCGTCAGAAAAGCCCTACTTTATGGATCAAGAATTTCTGAGTTTCGTA
AAAAGGAGCAACAAATTAAACAATGCAAGGGCACAGCAGGACAAGCACTCATTAAAGGAATGAGGA
GGGAACAGAACTCCAGGCCATTGTGAGATAGATACCCATTTAGGTATCTGTACCTGGAAACATT
TCCTTCTAAGAGCCATTACAGAATAGAAGATGAGACCACTAGAGAAAAGTTAGTGAATTTTTTT
TTAAAAGACCTAATAAACCCCTATTCTTCCTCAAAA

FIGURE 32

MSGRDTILGLCILALALSLAMMETFRFITTLVHIFISLVILGLLFVCGVLWWLYDYTDNLSIE
LDTERENMKCVLGFAIVSTGITAVLLVLI FVLRKRIKLTVELFQITNKAISSAPFLLFQPLWTF
ILIFFWVLWVAVLLSLGTAGAAQVMEGGQVEYKPLSGIRYMWSYHLIGLIWTSEFILACQQMTIA
GAVVTCYFNRSKNDDPPDHPILSSLSILFFYHQGTVVKGSFLISVVRIPRIIVMYMQNALKEQQHG
ALSRYLFRCCYCCFWCLDKYLLHLNQAYTTTAINGTDFCTSAKDAFKILSKNSSHFTSINCFGD
FIIIFLGKVLVVCFTVFGGLMAFNYNRAFQVWAVPLLLVAFFAYLVVHSFSLVFETVLDALFLCFA
VDLETNDGSSEKPYFMDQEFSLFVKRSNKLNNARAQQDKHSLRNEEGTELQAIVR

Important features:

Signal peptide:

amino acids 1-20

Putative transmembrane domains:

amino acids 35-54, 75-97, 126-146, 185-204, 333-350, 352-371

N-glycosylation sites.

amino acids 204-208, 295-299, 313-317

N-myristoylation sites.

amino acids 147-153, 178-184, 196-202, 296-275, 342-348

FIGURE 33

GTTCGATTAGCTCCTCTGAGAAGAAGAGAAAAGGTTCTTGGACCTCTCCCTGTTTCTTCTTAGA
 ATAATTTGTATGGGATTTGTGATGCAGGAAAGCCTAAGGGAAAAAGAATATTCATTCTGTGTGGT
 GAAAATTTTTTGAAGAAAAATTCGCTTCTTCAAACAAGGGTGTCTATTCTGATATTTATGAGGAC
 TGTGTCTCTACTATGAAGGCATCTGTTATGAAATGTTCTTGTGTTTGGTGGTGGTGGAGTAC
 ATTCAAACAAAGAAACGGCAAAGAAGATTAAAAGGCCCAAGTTCACTGTGCCTCAGATCAACTGC
 GATGTCAAAGCCGAAAGATCATCGATCCTGAGTTCATTGTGAAATGTCCAGCAGGATGCCAAGA
 CCCCAAATACCATGTTTATGGCACTGACGTGTATGCATCCTACTCCAGTGTGTGTGGCGCTGCCG
 TACACAGTGGTGTGCTTGATAATTCAAGAGGAAAAATCTTGTTCGGAAGGTTGCTGGACAGTCT
 GGTTACAAAGGGAGTTATTCCAACGGTGTCCAATCGTTATCCCTACCACGATGGAGAGAATCCTT
 TATCGTCTTAGAAAGTAAACCCAAAAAGGGTGTAACCTACCCATCAGCTCTTACATACTCATCAT
 CGAAAAGTCCAGTGCCTCAAGCAGGTGAGACCACAAAAGCCTATCAGAGGCCACCTATTCCAGGG
 ACAACTGCACAGCCGGTCACTCTGATGCAGCTTCTGGCTGTCACTGTAGCTGTGGCCACCCCCAC
 CACCTTGCCAAGGCCATCCCCTCTGCTGCTTCTACCACCAGCATCCCCAGACCACAATCAGTGG
 GCCACAGGAGCCAGGAGATGGATCTCTGGTCCACTGCCACCTACACAAGCAGCCAAACAGGCCC
 AGAGCTGATCCAGGTATCCAAAGGCAAGATCCTTCAGGAGCTGCCTTCCAGAAACCTGTTGGAGC
 GGATGTCAGCCTGGGACTTGTTCAAAAGAAGAATTGAGCACACAGTCTTGGAGCCAGTATCCC
 TGGGAGATCCAACTGCAAAATTGACTTGTCTGTTTTAATTGATGGGAGCACCAGCATTTGCCAAA
 CGGCGATTCCGAATCCAGAAGCAGCTCCTGGCTGATGTTGCCCAAGCTCTTGACATTGGCCCTGC
 CGGTCCACTGATGGGTGTTGTCCAGTATGGAGACAACCTGCTACTCACTTTAACCTCAAGACAC
 ACACGAATTCTCGAGATCTGAAGACAGCCATAGAGAAAATTACTCAGAGAGGAGGACTTTCTAAT
 GTAGGTCGGGCCATCTCCTTTGTGACCAAGAACTTCTTTTCAAAGCCAATGGAAACAGAAGCGG
 GGCTCCCAATGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT
 GACTTGGCAGAGAGTCAGGAATCAACATTTTCTTCATCACCATTGAAGGTGCTGCTGAAAATGAG
 AAGCATGATGTGGTGGAGCCCACTTTGCAAAACAAGGCCGTGTGCAGAACAACCGCTTCTACTC
 GCTCCACGTGCAGAGCTGGTTTGGCCTCCACAAGACCCTGCAGCCTCTGGTGAAGCGGGTCTGCG
 ACACTGACCGCTGGCCTGCAGCAAGACCTGCTTGAAGTGGCTGACATTGGCTTCGTCATCGAC
 GGCTCCAGCAGTGTGGGACGGGCAACTTCCGCACCGTCTCCAGTTTGTGACCAACCTCACCAA
 AGAGTTTGAGATTTCCGACACGGACACGCGCATCGGGGCCGTGCAGTACACCTACGAACAGCGGC
 TGGAGTTTGGGTTTCGACAAGTACAGCAGCAAGCCTGACATCCTCAACGCCATCAAGAGGGTGGGC
 TACTGGAGTGGTGGCACCAGCACGGGGGCTGCCATCAACTTCGCCCTGGAGCAGCTCTTCAAGAA
 GTCCAAGCCCAACAAGAGGAAGTTAATGATCCTCATCACCAGCGGAGGTCTTACGACGACGTCC
 GGATCCAGCCATGGCTGCCATCTGAAGGGAGTGATCACCTATGCGATAGGCGTTGCCCTGGGCT
 GCCCAAGAGGAGCTAGAAGTCATTGCCACTCACCCGCCAGAGACCACTCCTTCTTTGTGGACGA
 GTTTGACAACCTCCATCAGTATGTCCCAGGATCATCCAGAACATTTGTACAGAGTTCAACTCAC
 AGCCTCGGAACTGAATTCAAGAGCAGGCAGAGCACCAGCAAGTGTCTGCTTTACTAACTGACGTGTT
 GGACCACCCACCGCTTAATGGGGCACGCACGGTGCATCAAGTCTTGGGCAGGGCATGGAGAAAC
 AAATGTCTTGTATTATTTCTTGGCATCATGCTTTTTCATATTCCAAAACCTGGAGTTACAAAGA
 TGATCACAAACGTATAGAATGAGCCAAAAGGCTACATCATGTTGAGGGTGTGGAGATTTTACAT
 TTTGACAATTGTTTTCAAATAAATGTTGGAATACAGTGCAGCCCTTACGACAGGCTTACGTAG
 AGCTTTTGTGAGATTTTAAGTTGTTATTTCTGATTTGAACTCTGTAACCTCAGCAAGTTTCAT
 TTTTGTGATGACAATGTAGGAATTGCTGAATTAATGTTTAGAAGGATGAAAAATAAAAAA
 AA
 AA

FIGURE 34

MRTVVLTMKASVIEMFLVLLVTGVHSNKETAKKIKRPKFTVPQINCDVKAGKIIDPEFIVKCPAG
CQDPKYHVGTDVYASYSSVCGAAVHSGVLDNSGGKILVRKVAGQSGYKGSYNGVQSLSLPRWR
ESFIVLESKPKKGVTPYSALTYSSSKSPAAQAGETTKAYQRPPPIPGTTAQPVTLMQLLAVTVAVA
TPTTLPRPSPSAASTTSIPRPQSVGHRSEQEMDLWSTATYTSSQNRPRADPGIQRQDPGAAFAQP
VGADVSLGLVPKEELSTQSLEPVS LGDPNCKIDLSFLIDGSTSIGKRRFRIQKQLLADVAQALDI
GPAGPLMGVVQYGDNPATHTFNLTHTNSRDLKTAIEKITQRGGLSNVGRAISFVTKNFFSKANGN
RSGAPNVVVVMVDGWPTDKVEEASRLARESGINIFFITIEGAAENKQYVVEPNFANKAVCRTNG
FYSLHVQSWFGLHKTLLQPLVKRVCDDTLACSKTCLNSADIGFVIDGSSSVGTGNFRTVLQFVTN
LTKEFEISDTRIGAVQYTYEQRLDFGDKYSSKPDILNAIKRVGYWSGGTSTGAAINFALEQL
FKKSKPNKRKLMILITDGRSYDDVRIPAMAAHLKGVITYAIGVAWAAQEELEVIATHPARDHSFF
VDEFNLHQYVPRIIQNICTEFNSQPRN

Important features:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 181-200

N-glycosylation sites.

amino acids 390-394, 520-524

N-myristoylation sites.

amino acids 23-29, 93-99, 115-121, 262-268, 367-373, 389-395,
431-437, 466-472, 509-515, 570-576, 571-577, 575-581, 627-633

Amidation site.

amino acids 304-308

FIGURE 35

CCGAGCACAGGAGATTGCCTGCGTTTAGGAGGTGGCTGCGTTGTGGGAAAAGCTATCAAGGAAGAAATTGC
 CAAACCATGTCTTTTTTCTGTTTTCAGAGTAGTTCACAACAGATCTGAGTGTTTTAATTAAGCATGGAAT
 ACAGAAAACAACAAAAAATTAAGCTTTAATTTTCATCTGGAATTCACAGTTTTCTTAGCTCCCTGGACCC
 GGTGACCTGTTGGCTCTTCCCGCTGGCTGCTCTATCACGTGGTGTCTCCGACTACTCACCCGAGTGTA
 AAGAACCTTCGGCTCGCGTGCTTCTGAGCTGCTGTGGATGGCCTCGGCTCTCTGGACTGTCCTTCCGAGTA
 GGATGTCACCTAGATCCCTCAAATGGAGCCTCCTGCTGCTGCTCACTCCTGAGTTTCTTTGTGATGTGGTAC
 CTCAGCCTTCCCCACTACAATGTGATAGAACGCGTGAAGTGGATGTACTTCTATGAGTATGAGCCGATTTA
 CAGACAAGACTTTCACTTCACACTTCGAGAGCATTCAAAGTCTCTCATCAAATCCATTTCTGGTCATTCT
 TGGTGACCTCCCACCTTCAGATGTGAAAGCCAGGCAGGCCATTAGAGTTACTTGGGGTGAAAAAAGTCT
 TGGTGGGGATATGAGGTTCTTACATTTTTCTTATTAGGCCAAGAGGCTGAAAAGGAAGACAAAATGTTGGC
 ATTGTCCTTAGAGGATGAACACCTTCTTTATGGTGACATAATCCGACAAGATTTTTTAGACACATATAATA
 ACCTGACCTTGAAGAACATTATGGCATTCAAGGTGGGTAAGTGAAGTTTGGCCCAATGCCAAGTACGTAATG
 AAGACAGACACTGATGTTTTCATCAATACTGGCAATTTAGTGAAGTATCTTTTAAACCTAAACCACTCAGA
 GAAGTTTTTTCACAGGTTATCCTCTAATTGATAATTATTCCTATAGAGGATTTTACCAAAAAACCCATATTT
 CTTACCAGGAGTATCCTTTCAAGGTGTTCCCTCCATACTGCAGTGGGTGGGTTATATAATGTCCAGAGAT
 TTGGTGCCAAGGATCTATGAAATGATGGGTCACGTAAAACCCATCAAGTTTGAAGATGTTTATGTCGGGAT
 CTGTTTGAATTTATTAAGTGAACATTATTCAGAGACACAAATCTTTCTTTCTATATAGAATCC
 ATTTGGATGCTGTCAACTGAGACGTGTGATTGCAGCCCATGGCTTTTCTTCCAAGGAGATCATCACTTTT
 TGGCAGGTCATGCTAAGGAACACCACATGCCATTATTAAGTTCACATTCTACAAAAAGCCTAGAAGGACAG
 GATACCTTGTGAAAGTGTAAATAAAGTAGGTACTGTGGAATTCATGGGGAGGTCAAGTGTCTGGCTT
 AACTGAACTGAAACTCATGAAAAACCCAGACTGGAGACTGGAGGTTACACTTGTGATTTATAGTCAGG
 CCCTTCAAAGATGATATGTGGAGGAATTAATATAAAGGAATTGGAGGTTTTGCTAAAGAAATTAATAGG
 ACCAAACAATTTGGACATGTCATTCTGTAGACTAGAATTTCTTAAAGGGTGTTACTGAGTTATAAGCTCA
 CTAGGCTGTAAAAACAAAACATGTAGAGTTTATTTATTGAACAATGTAGTCACTTGAAGGTTTTGTGTA
 TATCTTATGTGGATTACCAATTTAAAAATATATGTAGTTCTGTGTCAAAAACTTCTTCACTGAAGTTATA
 CTGAACAAAATTTTACCTGTTTTTGGTCATTTATAAAGTACTTCAAGATGTTGCAGTATTTTACAGTTATT
 ATTATTTAAATTAAGTCACTTCACTTTGTGTTTTTAAATGTTTTGACGATTTCAATACAAGATAAAAGGATAG
 TGAATCATTTCTTACATGCAACATTTTCCAGTTACTTAACTGATCAGTTTATTATTGATACATCACTCCA
 TTAATGTAAAGTCATAGGTCATTATTGCATATCAGTAATCTCTTGGACTTTGTTAAATATTTTACTGTGGT
 AATATAGAGAAGAATTAAAGCAAGAAAATCTGAAAA

FIGURE 36

MASALWTVLPSRMSLRSLKWSLLLLSLLSFFVMWYLSLPHYNVIERVNWMYFYEYEPYRQDFHF
TLREHSNCSEQNPFLVILVTSHPSDVKARQAIRVTWGEKKSWWGYEVLTFLLGQEAEEKDKMLA
LSLEDEHLLYGDIIRQDFLDTYNNLTTLKTIMAFRWVTEFCPNAKYVMKTDTDVFINTGNLVKYL
NLNHSEKFFTGYPLIDNYSYRGFYQKTHISYQYEPFKVFPPYCSGLGYIMSRDLVPRIYEMMGHV
KPIKFEDVYVGICLNLLKVNHIHIPEDTNLFFLYRIHLDVCQLRRVIAAHGFSSKEIITFWQVMLR
NTTCHY

Important features:

Type II transmembrane domain:

amino acids 20-39

N-glycosylation sites.

amino acids 72-76, 154-158, 198-202, 212-216, 326-330

Glycosaminoglycan attachment site.

amino acids 239-243

Ly-6 / u-PAR domain proteins.

amino acids 23-37

N-myristoylation site.

amino acids 271-277

FIGURE 37

CGCTCGGGCACCAGCCGCGGGCAAGGATGGAGCTGGGTGCTGGACGAGTTGGGGCTCACTTTTCTTCAGCTCCTTCTCATC
TCGTCTTGGCCAGAGAGTACACAGTCATTAAATGAAGCCTGCCCTGGAGCAGAGTGGAAATATCATGTGTCGGGAGTGCTGTG
AATATGATCAGATTGAGTGCCTCTGCCCCGGAAGAGGGAGTCTGGGTTATACCATCCCTTGCTGCAGGAATGAGGAGAA
TGAGTGTGACTCCTGCCTGATCCACCCAGGTTGTACCATCTTGAAAAGCTGCAAGAGCTGCCGAAATGGCTCATGGGGGGGT
ACCTTGGATGACTTCTATGTGAAGGGTTCTACTGTGCAGAGTGCCGAGCAGGCTGGTACGGAGGAGACTGCATGCGATGTG
GCCAGGTTCTGCGAGCCCCAAAGGGTCAGATTTTGTGGAAAGCTATCCCTAAATGCTCACTGTGAATGGACCATTTCATGC
TAAACCTGGGTTTGTTCATCCAACTAAGATTTGTTCATGTTGAGTCTGGAGTTTGACTACATGTGCCAGTATGACTATGTTGAG
GTTCTGTATGGAGACAACCGCATGGCCAGATCATCAAGCGTGTCTGTGGCAACGAGCGGCCAGCTCCTATCCAGAGCATAG
GATCCTCACTCCAGTCCCTCTTCCACTCCGATGGCTCCAAGAAATTTGACGGTTTCCATGCCATTTATAGGAGATCACAGC
ATGCTCCTCATCCCCTTGTTTCCATGACGCGACGTGCGTCTTGACAAGGCTGGATCTTACAAGTGTGCTGCTTGGCAGGC
TATACTGGGCAGCGCTGTGAAAATCTCCTTGAAGAAAGAACTGCTCAGACCTGGGGGCCAGTCAATGGGTACAGAAAA
TAACAGGGGGCCCTGGGCTTATCAACGGACGCCATGCTAAAATTGGCACCCTGGTGTCTTCTTTTGTAACTCCTATGT
TCTTAGTGGCAATGAGAAAAGAACTTGCAGCAGAAATGGAGAGTGGTCAGGGAAACAGCCCATCTGCATAAAAGCCTGCCGA
GAACCAAAGATTTAGACCTGGTGAGAAGGAGGTTCTTCCGATGCAGTTTCAGTCAAGGGAGACACCATTACACAGCTAT
ACTCAGCGCCCTTCAGCAAGCAGAACTGCAGAGTCCCCCTACCAAGAACCCAGCCCTTCCCTTTGGAGATCTGCCCATGGG
ATACCAACATCTGCATACCCAGCTCCAGTATGAGTGCATCTCACCCTTCTACCGCGCCTGGGCAGCAGCAGGAGGACATGT
CTGAGGACTGGGAAGTGGAGTGGCGGGCACCATCTGCATCCCTATCTGCGGGAAAATTGAGAACATCACTGCTCCAAAGA
CCCAAGGGTTGCGCTGGCCGTGGCAGGCAGCCATCTACAGGAGGACCAGCGGGTGCATGACGGCAGCTACACAAGGGAGC
GTGGTTCTTAGTCTGCAGCGGTGCCCTGGTGAATGAGCGCACTGTGGTGGTGGCTGCCACTGTGTACTGACCTGGGGAAG
GTCACCATGATCAAGACAGCAGACCTGAAAGTTGTTTGGGGAAATCTACCGGGATGATGACCGGGATGAGAAGACCATCC
AGAGCCTACAGATTTCTGCTATCATTTGCATCCCAACTATGACCCCATCTGCTTGTATGCTGACATCGCCATCCTGAAGCT
CCTAGACAAGGCCGTATCAGCACCAGTCCAGCCCATCTGCCTCGCTGCCAGTCCGGATCTCAGCACTTCTTCCAGGAG
TCCCACATCACTGTGGCTGGCTGGAATGTCTTGGCAGAGCTGAGGAGCCCTGGCTTCAAGAACGACACACTGCGCTCTGGGG
TGGTCAGTGTGGTGGACTCGCTGCTGTGTGAGGAGCAGCATGAGGACCATGGCATCCAGTGAGTGTCACTGATAACATGTT
CTGTGCCAGCTGGGAACCCACTGCCCTTCTGATATCTGCATGCAAGAGACAGGAGGCATCGCGGCTGTGTCTTCCCGGA
CGAGCATCTCCTGAGCCACGCTGGCATCTGATGGACTGCTCAGCTGGAGCTATGATAAAACATGCAGCCACAGGCTCTCCA
CTGCCCTTACCAAGGTGCTGCCCTTTTAAAGACTGGATTGAAAGAAATATGAAATGAACCATGCTCATGCACTCCTTGAGAAG
TGTTTCTGTATATCCGTCTGTACGTGTGTCATTGCGTGAAGCAGTGTGGCCTGAAGTGTGATTGGCTGTGAAGTTGGCT
GTGCCAGGGCTTCTGACTTCAGGGACAAAACCTCAGTGAAGGGTGAAGTAGACCTCCATGCTGGTAGGCTGATGCCGCTCCA
CTACTAGGACAGCCAATTGGAAGATGCCAGGGCTTGAAGAAGTAAGTTTCTTCAAGAAGACCATATACAAAACCTCTCCA
CTCCACTGACCTGGTGGCTTCCCCAACTTTCAGTTATACGAATGCCATCAGCTTGACCAGGGAAGATCTGGGCTTCATGAG
GCCCCCTTTGAGGCTCTCAAGTTCTAGAGAGCTGCCTGTGGGACAGCCAGGGCAGCAGAGCTGGGATGTGGTGCATGCCCTT
TGTTGATACATGGCCACAGTACAGTCTGGTCTTTTCTTCCCCATCTCTGTACACATTTTAATAAAATAAGGGTTGGCTTCT
GAAC TACAAA
AAA

FIGURE 38

MELGCWTQLGLTFLQLLLISSLPREYTVINEACPGAENIMCRECCEYDQIECVCPGKREVVGYT
IPCCRNEENECDSCLIHPGCTIFENCKSCRNGSWGGLDDFYVKGFYCAECRAGWYGGDCMRGQ
VLRAPKGQILLESYPLNAHCEWTIHAKPGFVIQLRFVMLSLEFDYMCQYDYVEVRDGDNRDQII
KRVCGNERPAPIQSIGSSHLVLFHSDGSKNFDGFHAIYEEITACSSSPCFHDGTCVLDKAGSYKC
ACLAGYTGQRCENLLEERNCSDPGGPVNGYQKITGGPGLINGRHAKIGTVVSFFCNSYVLSGNE
KRTCQONGEWSGKQPICIKACREPKISDLVRRRVLPQVQSRETPHQLYSAAFSKQKLQSAPTK
KPALPFGDLPMGYQHLHTQLQYECISPFYRRLGSSRRTCLRTGKWSGRAPSCIPICGKIENITAP
KTQGLRWPQAAIYRRTSGVHDGSLHKGAWFLVCSGALVNERTVVVAACHCVTDLGKVTMIKTADL
KVVLGKFYRDDDRDEKTIQSLQISAILHPNYDPILLDADIAILKLLDKARISTRVQPICLAASR
DLSTSFQESHITVAGWNVLADVRSPGFKNDTLRSGVSVVDSLLCEEQHEDHGIPVSVTDNMFCA
SWEPTAPSDICTAETGGIAAVSFPGRASPEPRWHLMLVSWSYDKTCSHRLSTAFTKVLPFKDWI
ERNMK

Important features of the protein:

Signal peptide:

amino acids 1-23

EGF-like domain cysteine pattern signature.

amino acids 260-272

N-glycosylation sites.

amino acids 96-100, 279-283, 316-320, 451-455, 614-618

N-myristoylation sites.

amino acids 35-41, 97-103, 256-262, 284-290, 298-304, 308-314,
474-480, 491-497, 638-644, 666-672

Amidation site.

amino acids 56-60

Serine proteases, trypsin family.

amino acids 489-506

CUB domain proteins profile.

amino acids 150-167

FIGURE 39

GGTTCCTACATCCTCTCATCTGAGAATCAGAGAGCATAATCTTCTTACGGGCCCCGTGATTTATTAACGTGGCTTAATC
TGAAGGTTCTCAGTCAAATCTTTGTGATCTACTGATTGTGGGGGCATGGCAAGGTTTGCTTAAAGGAGCTTGGCTGG
TTTGGGCCCTTGTAGCTGACAGAAGGTGGCCAGGGAGAATGCAGCACACTGCTCGGAGAAATGAAGGCGCTTCTGTTGC
TGGTCTTGCCTTGGCTCAGTCTGCTAACTACATTGACAATGTGGGCAACCTGCACTTCCTGTATTGAGAACTCTGTA
AAGGTGCCTCCCACTACGGCCTGACCAAAGATAGGAAGAGGCGCTCACAAGATGGCTGTCCAGACGGCTGTGCGAGCC
TCACAGCCACGGCTCCCTCCCGAGAGGTTTCTGCAGCTGCCACCATCTCCTTAATGACAGACGAGCCTGGCCTAGACA
ACCCCTGCCTACGTGTCTCGGCAGAGGACGGGCAGCCAGCAATCAGCCAGTGGACTCTGGCCGGAGCAACCGAACTA
GGGCACGGCCCTTTGAGAGATCCACTATTAGAAGCAGATCATTTAAAAAATAAATCGAGCTTTGAGTGTTCTTCGAA
GGACAAAGAGCGGGAGTGCAGTTGCCAACCATGCCGACCAGGGCAGGAAAAATTCTGAAACACCACTGCCCTGAAG
TCTTTCCAAGGTTGTACCACCTGATTCCAGATGGTGAAATTACCAGCATCAAGATCAATCGAGTAGATCCCAGTGAA
GCCTCTCTATTAGGCTGGTGGGAGGTAGCGAAACCCCACTGGTCCATATCATTATCCAACACATTTATCGTGATGGG
TGATCGCCAGAGACGGCCGGCTACTGCCAGGAGACATCATTCTAAAGGTCAACGGGATGGACATCAGCAATGTCCCTC
ACAACTACGCTGTGCGTCTCTGCGGCAGCCCTGCCAGGTGCTGTGGCTGACTGTGATGCGTGAACAGAAGTCCGCA
GCAGGAACAATGGACAGGCCCCGGATGCCACAGACCCGAGATGACAGCTTTTCATGTGATTCTCAACAAAAGTAGCC
CCGAGGAGCAGCTTGAATAAACTGGTGCAGAGGTGGATGAGCCTGGGGTTTTTCATCTTCAATGTGCTGGATGGCG
GTGTGGCATATCGACATGGTCAGCTTGGAGAGAATGACCGTGTGTAGCCATCAATGGACATGATCTTCGATATGGCA
GCCCAGAAAGTGGGCTCATCTGATTGAGGCCAGTGAAGACGTGTTACCTCGTCGTGTCCCGCCAGGTTTCGGCAGC
GGAGCCCTGACATCTTTAGGAAGCCGGCTGGAACAGCAATGGCAGCTGGTCCCCAGGGCCAGGGGAGAGGAGCAACA
CTCCCAAGCCCTCCATCTACAATTACTTGTGATGAGAAGGTGGTAAATATCCAAAAGACCCCGTGAATCTCTCG
GCATGACCGTCGAGGGGGAGCATCACATAGAGAATGGGATTGCTTATCTATGTCATCAGTGTGAGCCCGAGGAG
TCATAAGCAGAGATGGAAGAATAAAACAGGTGACATTTTGTGAATGTGGATGGGGTCGAACTGACAGAGTCAAGC
GGAGTGAGGCAGTGGCATTATTGAAAAGAACATCATCTCGATAGTACTCAAAGCTTTGGAAGTCAAAGATATGAGC
CCCAGGAAGACTGCAGCAGCCAGCAGCCCTGGACTCCAACCACAACATGGCCCCACCCAGTACTGGTCCCCATCCT
GGGTGATGTGGCTGGAATTACCACGGTGCTTGTATACTGTAAAGATATTGTATTACGAAGAAACACAGCTGGAAGTC
TGGGCTTCTGCATTGTAGGAGGTTATGAAGAATACAATGGAACAAACCTTTTTTCATCAAATCCATTGTTGAAGGAA
CACCAGCATACAATGATGGAAGAATTAGATGTGGTGATATTCTTCTTGTGTCATGGTAGAAGTACATCAGGAATGA
TACATGCTTGCTTGGCAAGACTGCTGAAAGAACTTAAAGGAAGAATTACTCTAACTATTGTTTCTTGGCCTGGCACTT
TTTTATAGAATCAATGATGGGTGAGGAAAAACAGAAAATCACAAATAGGCTAAGAAGTTGAAACACTATATTTATC
TTGTCAGTTTTTATATTTAAAGAAAGAAATACATTGTAAAAATGTCAGGAAAAGTATGATCATCTAATGAAAGCCAGTT
ACACCTCAGAAAATATGATTTCAAAAAAATTAAGTACTAGTTTTTTTTTTCAGTGTGGAGGATTTCTCATTACTCTAC
AACATTGTTTATATTTTTCTATTCAATAAAAAAGCCCTAAAAACAATAAATGATTGATTGTATACCCCACTGAATT
CAAGCTGATTTAAATTTAAATTTGGTATATGCTGAAGTCTGCCAAGGGTACATTATGGCCATTTTAAATTTACAGCT
AAAAATTTTTTAAATGCATTGCTGAGAAACGTTGCTTTCATCAAACAAGAATAAATATTTTTCAGAAGTTAA

FIGURE 40

MKALLLLVLPWLS PANYIDNVGNLHFLYSELCKGASHYGLTKDRKRRSQDGCPCDGCASLTATAPS
PEVSAAATISLMTDEPGLDNPAYVSSAEDGQPAISPVDSGRSNRTRARPFERSTIRSRSFKKINR
ALSVLRRTKSGSAVANHADQGRESENTTAPEVFPRLYHLIPDGEITSIKINRVDPSSELSIRLV
GGSETPLVHIIIQHIYRDGVIARDGRLLPGDIILKVNGMDISNVPHNYAVRLLRQPCQVLWLTVM
REQKFRSRNNGQAPDAYRPRDDSFHVILNKSSPEEQLGIKLVRKVDEPGVFIFNVLDGGVAYRHG
QLEENDRVLAINGHDLRYGSPESAHLIQASERRVHLVVSQRQVRQSPDIFQEAGWNSNGSWSPG
PGRSNTPKPLHPTITCHEKVNIQKDPGESLGMTVAGGASHREWDLPIYVISVEPGGVISRDR
IKTGDILLNVGVELTEVSRSEAVALLKRTSSSIVLKALEVKEYEPQEDCSSPAALDSNHNMAP
SDWSPSWVMWLELPCLYNCKDIVLRRNTAGSLGFCIVGGYEEYNGNKPFFIKSIVEGTPAYNDG
RIRCGDILLAVNGRSTSGMIHACLARLLKELKGRITLTIVSWPGTFL

Important features:

Signal peptide:

amino acids 1-15

N-glycosylation sites.

amino acids 108-112, 157-161, 289-293, 384-388

Tyrosine kinase phosphorylation sites.

amino acids 433-441, 492-500

N-myristoylation sites.

amino acids 51-57, 141-147, 233-239, 344-350, 423-429, 447-453,
467-473, 603-609

FIGURE 41

ACCAGGCATTGTATCTTCAGTTGTCATCAAGTTCGCAATCAGATTGGAAAAGCTCAACTTGAAGCTTT
 CTTGCCTGCAGTGAAGCAGAGAGATAGATATTATTACGTAATAAAAAACATGGGCTTCAACCTGACT
 TTCCACCTTTCTACAAATTCGATTACTGTTGCTGTTGACTTTGTGCCTGACAGTGTTGGTGGGC
 CACCAGTAACCTTCGTGGGTGCCATTCAAGAGATTCTTAAAGCAAAGGAGTTCATGGCTAATTTCC
 ATAAGACCCCTCATTTTGGGGAAGGGAAGAACTCTGACTAATGAAGCATCCACGAAGAAGGTAGAACTT
 GACAACTGTCCTTCTGTGTCTCCTTACCTCAGAGGCCAGAGCAAGCTCATTTCACACAGATCTCAC
 TTTGGAAGAGGTACAGGCAGAAAATCCCAAAGTGTCCAGAGGCCGATCGCCCTCAGGAATGTAAAG
 CTTTACAGAGGGTCGCCATCTCCTTCCCCACCGAACAGAGAGAAACACCTGATGTACCTGCTGGAA
 CATCTGCATCCCTTCCTGCAGAGGCAGCAGCTGGATTATGGCATCTACGTCATCCACAGGCTGAAGG
 TAAAAAGTTTAAATCGAGCCAACTCTTGAATGTGGGCTATCTAGAAGCCCTCAAGGAAGAAAATTGGG
 ACTGCTTTATATTCCACGATGTGGACCTGGTACCCGAGAATGACTTTAACCTTTACAAGTGTGAGGAG
 CATCCCAAGCATCTGGTGGTTGGCAGGAACAGCAGCTGGGTACAGGTTACGTTACAGTGGATATTTGG
 GGGTGTACTGCCCTAAGCAGAGAGCAGTTTTTCAAGGTGAATGGATTCTCTAACAACCTACTGGGGAT
 GGGGAGGCGAAGACGATGACCTCAGACTCAGGGTTGAGCTCCAAAGAATGAAAATTTCCCGGCCCTG
 CCTGAAGTGGGTAAATATACAATGGTCTTCCACACTAGAGACAAAGGCAATGAGGTGAACGCAGAACG
 GATGAAGCTCTTACACCAAGTGTACGAGTCTGGAGAACAGATGGGTTGAGTAGTTGTTCTTATAAAT
 TAGTATCTGTGGAACACAATCCTTTATATATCAACATCACAGTGGATTTCTGGTTTGGTGCATGACCC
 TGGATCTTTTGGTGATGTTTGAAGAACTGATTCTTTGTTTGCAATAATTTTGGCCTAGAGACTTCAA
 ATAGTAGCACACATTAAGAACCTGTTACAGCTCATTGTTGAGCTGAATTTTCCTTTTGTATTTCT
 TAGCAGAGCTCCTGGTGATGTAGAGTATAAAACAGTTGTAACAAGACAGCTTCTTAGTCATTTTGAT
 CATGAGGGTTAAATATTGTAATATGGATACTTGAAGGACTTTATATAAAAGGATGACTCAAAGGATAA
 AATGAACGCTATTTGAGGACTCTGGTTGAAGGAGATTTATTTAAATTTGAAGTAATATATTATGGGAT
 AAAAGGCCACAGGAAATAAGACTGCTGAATGTCTGAGAGAACCAGAGTTGTTCTCGTCCAAGGTAGAA
 AGGTACGAAGATACAATACTGTTATTCATTTATCCTGTACAATCATCTGTGAAGTGGTGGTGTGAGGT
 GAGAAGGCGTCCACAAAAGAGGGGAGAAAAGGCGACGAATCAGGACACAGTGAAGTTGGGAATGAAGA
 GGTAGCAGGAGGGTGGAGTGTGCGCTGCAAAGGCAGCAGTAGCTGAGCTGGTTGCAGGTGCTGATAGC
 CTTTCAAGGGAGGACCTGCCCAGGTATGCCTTCCAGTGTGATGCCACCAGAGAATACATTCTCTATTAGT
 TTTTAAAGAGTTTTTGTAAAATGATTTTGTACAAGTAGGATATGAATTAGCAGTTTACAAGTTTACAT
 ATTAACATAATAAATATGTCTATCAAATACCTCTGTAGTAAATGTGAAAAGCAAAA

FIGURE 42

MGFNLTFHLSYKFRLLLLLTLCLTVVGWATSNYFVGAIQEIPKAKEFMANFHKTLLLGKGTLTN
EASTKKVELDNCPSVSPYLRGQSKLIFKPDLTLEEVQAENPKVSRGRYRPQECKALQRVAILVPH
RNREKHLMYLLEHLHPFLQRQQLDYGIYVIHQAEKKFNRAKLLNVGYLEALKEENWDCFI FHDV
DLVPENDFNLYKCEEHPKHLVVGGRNSTGYRLRYSYFGGVTALSREQFFKVNGFSNNYWGWGGE
DDLRLRVELQRMKISRPLPEVGKYTEMFHTTRDKGNEVNAERMKLLHQVSRVWRTDGLSSCSYKLV
SVEHNPLYINITVDFWFGA

Important features:

Signal peptide:

amino acids 1-27

N-glycosylation sites.

amino acids 4-8, 220-224, 335-339

Xylose isomerase proteins.

amino acids 191-202

FIGURE 43

GCTCAAGACCCAGCAGTGGGACAGCCAGACAGACGGCACGATGGCACTGAGCTCCCAGATCTGGG
CCGCTTGCCTCCTGCTCCTCCTCCTCGCCAGCCTGACCAGTGGCTCTGTTTTCCACAAACAG
ACGGGACAACCTGCAGAGCTGCAACCCCAGGACAGAGCTGGAGCCAGGGCCAGCTGGATGCCCAT
GTTCCAGAGGCGAAGGAGGCGAGACACCCACTTCCCCATCTGCATTTTCTGCTGCGGCTGCTGTC
ATCGATCAAAGTGTGGGATGTGCTGCAAGACGTAGAACCTACCTGCCCTGCCCCGTCCCCTCCC
TTCCTTATTTATTCCTGCTGCCCCAGAACATAGGTCTTGGAATAAAATGGCTGGTTCTTTTGTTT
TCCAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 44

MALSSQIWAACLLLLLLSLTSGSVFPQQTGQLAELQPQDRAGARASWMPMFQRRRRRDTHFPI
CIFCCGCCCHRSKCGMCKT

Important features:

Signal peptide:

amino acids 1-24

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 58-59

N-myristoylation site.

amino acids 44-50

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 1-12

FIGURE 45

GTGGCTTCATTTTCAGTGGCTGACTTCCAGAGAGCAATATGGCTGGTTCCCCAACATGCCTCACCC
TCATCTATATCCTTTGGCAGCTCACAGGGTCAGCAGCCTCTGGACCCGTGAAAGAGCTGGTCGGT
TCCGTTGGTGGGGCCGTGACTTTCCCCCTGAAGTCCAAAGTAAAGCAAGTTGACTCTATTGTCTG
GACCTTCAACACAACCCCTCTTGTCACCATACAGCCAGAAGGGGGCACTATCATAGTGACCCAAA
ATCGTAATAGGGAGAGAGTAGACTTCCCAGATGGAGGCTACTCCCTGAAGCTCAGCAAACCTGAAG
AAGAATGACTCAGGGATCTACTATGTGGGGATATACAGCTCATCACTCCAGCAGCCCTCCACCCA
GGAGTACGTGCTGCATGTCTACGAGCACCTGTCAAAGCCTAAAGTCACCATGGGTCTGCAGAGCA
ATAAGAATGGCACCTGTGTGACCAATCTGACATGCTGCATGGAACATGGGGAAGAGGATGTGATT
TATACCTGGAAGGCCCTGGGGCAAGCAGCCAATGAGTCCCATAATGGGTCCATCCTCCCCATCTC
CTGGAGATGGGGAGAAAGTGATATGACCTTCATCTGCGTTGCCAGGAACCCCTGTCAGCAGAACT
TCTCAAGCCCCATCCTTGCCAGGAAGCTCTGTGAAGGTGCTGCTGATGACCCAGATTCTCCATG
GTCTCTCTGTGTCTCCTGTGTGGTGCCCCCTCCTGCTCAGTCTCTTTGTACTGGGGCTATTTCTTTG
GTTTCTGAAGAGAGAGAGACAAGAAGAGTACATTGAAGAGAAGAAGAGAGTGGACATTTGTGGG
AAACTCCTAACATATGCCCCCATCTGGAGAGAACACAGAGTACGACACAATCCCTCACACTAAT
AGAACAATCCTAAAGGAAGATCCAGCAAATACGGTTTACTCCACTGTGGAAATACCGAAAAAGAT
GGAAAATCCCCACTCACTGCTCACGATGCCAGACACACCAAGGCTATTTGCCTATGAGAATGTTA
TCTAGACAGCAGTGCACTCCCCTAAGTCTCTGCTCA

FIGURE 46

MAGSPTCLTLIYILWQLTGSAAASGPVKELVGSVGGAVTFPLKSKVKQVDSIVWTFNTTPLVTIQP
EGGTIIIVTQNRNRERVDFPDGGYSLKLSKLKKNDSGIYYVGIYSSSLQQPSTQEYVLHVYEHLSK
PKVTMGLQSNKNGTCVTNLTCMEHGEEVDIYTWKALGQAANESHNGSILPISWRWGESDMTFIC
VARNPVSRNFSSPILARKLCEGAADDPDSSMVLLCLLLVPLLLSLFVLGLFLWFLKRERQEEYIE
EKKRVDICRETPNICPHSGENTYDTIPHTNRTILKEDPANTVYSTVEIPKKMENPHSLLTMPDT
PRLFAYENVI

Important features:

Signal peptide:

amino acids 1-22

Transmembrane domain:

amino acids 224-250

Leucine zipper pattern.

amino acids 229-251

N-glycosylation sites.

amino acids 98-102, 142-146, 148-152, 172-176, 176-180, 204-208,
291-295

FIGURE 47

GGCTCGAGCGTTTCTGAGCCAGGGTGACCAATGACCTGCTGCGAAGGATGGACATCCTGCAATGG
ATTCAGCCTGCTGGTTCTACTGCTGTTAGGAGTAGTTCTCAATGCGATACCTCTAATTGTCAGCT
TAGTTGAGGAAGACCAATTTCTCAAAACCCCATCTCTTGCTTTGAGTGGTGGTCCCAGGAATT
ATAGGAGCAGGTCTGATGGCCATTCCAGCAACAACAATGTCCTTGACAGCAAGAAAAAGAGCGTG
CTGCAACAACAGAACTGGAATGTTTCTTTCATCATTTTTCAGTGTGATCACAGTCATTGGTGCTC
TGTATTGCATGCTGATATCCATCCAGGCTCTCTTAAAAGGTCCTCTCATGTGTAATTCTCCAAGC
AACAGTAATGCCAATTGTGAATTTTCATTGAAAAACATCAGTGACATTCATCCAGAATCCTTCAA
CTTGCAAGTGGTTTTTCATGACTCTTGTGCACCTCCTACTGGTTTCAATAAACCCACCAGTAACG
ACACCATGGCGAGTGGCTGGAGAGCATCTAGTTTCCACTTCGATTCTGAAGAAAACAAACATAGG
CTTATCCACTTCTCAGTATTTTtaggtctattgcttgttggaattctggaggtcctgcttgggct
CAGTCAGATAGTCATCGGTTTCCTTGGCTGTCTGTGTGGAGTCTCTAAGCGAAGAAGTCAAATTG
TGTAGTTTAATGGGAATAAAATGTAAGTATCAGTAGTTTGAAAAAAAAA

FIGURE 48

MTCCEGWTS CNGFSLLVLLLLGVVLNAIPLIVSLVEEDQFSQNPISCFEWWFPGIIGAGLMAIPA
TTMSLTARKRACCNRTGMFLSSFFSVITVIGALYCLISIQALLKGPLMCNSPSNSNANCEFSL
KNISDIHPESFNLQWFFNDSCAPPTGFNKPTSNDTMASGWRASSFHFDEENKHRLIHFSVFLGL
LLVGILEVLFGLSQIVIGFLGCLCGVSKRRSQIV

Important features:

Transmembrane domains:

amino acids 10-31 (type II), 50-72, 87-110, 191-213

N-glycosylation sites.

amino acids 80-84, 132-136, 148-152, 163-167

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 223-227

N-myristoylation sites.

amino acids 22-28, 54-60, 83-89, 97-103, 216-222

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 207-218

TNFR/NGFR family cysteine-rich region protein.

amino acids 4-12

FIGURE 49

ATCCGTTCTCTGCGCTGCCAGCTCAGGTGAGCCCTCGCCAAGGTGACCTCGCAGGACACTGGTGA
AGGAGCAGTGAGGAACCTGCAGAGTCACACAGTTGCTGACCAATTGAGCTGTGAGCCTGGAGCAG
ATCCGTGGGCTGCAGACCCCGCCCCAGTGCCTCTCCCCCTGCAGCCCTGCCCCTCGAACTGTGA
CATGGGAGAGAGTGACCCTGGCCCTTCTCCTACTGGCAGGCCTGACTGCCTTGGAGCCAATGACC
CATTTGCCAATAAAGACGATCCCTTCTACTATGACTGGAAAAACCTGCAGCTGAGCGGACTGATC
TGCGGAGGGCTCCTGGCCATTGCTGGGATCGCGGCAGTTCTGAGTGGCAAATGCAAATACAAGAG
CAGCCAGAAGCAGCACAGTCCTGTACCTGAGAAGGCCATCCCACTCATCACTCCAGGCTCTGCCA
CTACTTGCTTGAGCACAGGACTGGCCTCCAGGGATGGCCTGAAGCCTAACACTGGCCCCCAGCACC
TCCTCCCCCTGGGAGGCCTTATCCTCAAGGAAGGACTTCTCTCCAAGGGCAGGCTGTTAGGCCCCCT
TTCTGATCAGGAGGCTTCTTTATGAATTAACTCGCCCCACCACCCCTCA

FIGURE 50

MERVTLALLLLAGLTALEANDPFANKDDPFYYDWKNLQLSGLICGGLLAIAGIAAVLSGKCKYKS
SQKQHSPVPEKAIPILITPGSATTC

Important features:

Signal peptide:

amino acids 1-16

Transmembrane domain:

amino acids 36-59

N-myristoylation sites.

amino acids 41-47, 45-51, 84-90

Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7.

amino acids 54-67

FIGURE 51

GTGGACTCTGAGAAGCCCAGGCAGTTGAGGACAGGAGAGAGAAGGCTGCAGACCCAGAGGGAGGG
AGGACAGGGAGTCGGAAGGAGGAGGACAGAGGAGGGCACAGAGACGCAGAGCAAGGGCGGCAAGG
AGGAGACCCTGGTGGGAGGAAGACACTCTGGAGAGAGAGGGGGCTGGGCAGAGATGAAGTTCCAG
GGGCCCCCTGGCCTGCCTCCTGCTGGCCCTCTGCCTGGGCAGTGGGGAGGCTGGCCCCCTGCAGAG
CGGAGAGGAAAGCACTGGGACAAATATTGGGGAGGCCCTTGGACATGGCCTGGGAGACGCCCTGA
GCGAAGGGGTGGGAAAGGCCATTGGCAAAGAGGCCGAGGGGCAGCTGGCTCTAAAGTCAGTGAG
GCCCTTGGCCAAGGGACCAGAGAAGCAGTTGGCACTGGAGTCAGGCAGGTTCCAGGCTTTGGCGC
AGCAGATGCTTTGGGCAACAGGCTCGGGGAAGCAGCCCATGCTCTGGGAAACACTGGGCACGAGA
TTGGCAGACAGGCAGAAGATGTCATTGACACGGAGCAGATGCTGTCCGCGGCTCCTGGCAGGGG
GTGCCTGGCCACAGTGGTGCTTGGGAACTTCTGGAGGCCATGGCATCTTTGGCTCTCAAGGTGG
CCTTGGAGGCCAGGGCCAGGGCAATCCTGGAGGTCTGGGGACTCCGTGGGTCCACGGATACCCCCG
GAAACTCAGCAGGCAGCTTTGGAATGAATCCTCAGGCAGCTCCCTGGGGTCAAGGAGGCAATGGA
GGGCCACCAAACCTTTGGGACCAACACTCAGGGAGCTGTGGCCAGCCTGGCTATGGTTCAGTGAG
AGCCAGCAACCAGAATGAAGGTCACGAATCCCCCACCATCTGGCTCAGGTGGAGGCTCCAGCA
ACTCTGGGGGAGGCAGCGGCTCACAGTCGGGCAGCAGTGGCAGTGGCAGCAATGGTGACAACAAC
AATGGCAGCAGCAGTGGTGGCAGCAGCAGTGGCAGCAGCAGTGGCAGCAGCAGTGGCGGCAGCAG
TGGCGGCAGCAGTGGTGGCAGCAGTGGCAACAGTGGTGGCAGCAGAGGTGACAGCGGCAGTGAGT
CCTCCTGGGGATCCAGCACCGGCTCCTCCTCCGGCAACCACGGTGGGAGCGGCGGAGGAAATGGA
CATAAACCCGGGTGTGAAAAGCCAGGGAATGAAGCCCGCGGGAGCGGGGAATCTGGGATTCAGGG
CTTCAGAGGACAGGGAGTTTCCAGCAACATGAGGGAATAAGCAAAGAGGGCAATCGCCTCCTTG
GAGGCTCTGGAGACAATTATCGGGGGCAAGGTCGAGCTGGGGCAGTGGAGGAGGTGACGCTGTT
GGTGGAGTCAATACTGTGAACCTCTGAGACGTCTCCTGGGATGTTTAACTTTGACACTTTCTGGAA
GAATTTTAAATCCAAGCTGGGTTTCATCAACTGGGATGCCATAAACAAGGACCAGAGAAGCTCTC
GCATCCCGTGAACCTCCAGACAAGGAGCCACCAGATTGGATGGGAGCCCCCACACTCCCTCCTTAA
AACACCACCCTCTCATCTAATCTCAGCCCTTGCCCTGAAATAAACCTTAGCTGCCCCACAAA
AA
AA

FIGURE 52

MKFQGPLACLLALCLGSGEAGPLQSGEESTGTNIGEALGHGLGDALSEGVGKAIGKEAGGAAGSKVS
EALQGQGTREAVGTGVRQVPFGAADALGNRVGEAAHALGNTGHEIGRQAEDVIRHGADAVRGSWQGV
GHSGAWETSGGHGIFGSQGGGQGGQGNPGLGTPWVHGYPGNSAGSFGMNPQGAPWGQGGNGGPPNF
GTNTQGAQAQPGYGSVRASNQNEGCTNPPPSGSGGGSSNSGGGSGSQSGSSGSGSNGDNNNGSSSSGGS
SSGSSSSGSSSGSSSGSSSGSSGSGSGSRGDSGSESSWGSSTGSSSGNHGGSGGGNGHKPGCEKPGNE
ARGSGESGIQGFRRQGVSSNMREISKEGNRLGSGDNYRGQSSWGSGGGDAVGGVNTVNSETPGM
FNFDTFWKNFKSKLGFINWDAINKDQRSSRIP

Signal peptide:

amino acids 1-21

N-glycosylation site.

amino acids 265-269

Glycosaminoglycan attachment site.

amino acids 235-239, 237-241, 244-248, 255-259, 324-328, 388-392

Casein kinase II phosphorylation site.

amino acids 26-30, 109-113, 259-263, 300-304, 304-308

N-myristoylation site.

amino acids 17-23, 32-38, 42-48, 50-56, 60-66, 61-67, 64-70, 74-80,
90-96, 96-102, 130-136, 140-146, 149-155, 152-158, 155-161,
159-165, 163-169, 178-184, 190-196, 194-200, 199-205, 218-224,
236-242, 238-244, 239-245, 240-246, 245-251, 246-252, 249-252,
253-259, 256-262, 266-272, 270-276, 271-277, 275-281, 279-285,
283-289, 284-290, 287-293, 288-294, 291-297, 292-298, 295-301,
298-304, 305-311, 311-317, 315-321, 319-325, 322-328, 323-329,
325-331, 343-349, 354-360, 356-362, 374-380, 381-387, 383-389,
387-393, 389-395, 395-401

Cell attachment sequence.

amino acids 301-304

FIGURE 53

GGAGAAGAGGTTGTGTGGGACAAGCTGCTCCCGACAGAAGGATGTCGCTGCTGAGCCTGCCCTGG
CTGGGCCTCAGACCGGTGGCAATGTCCCATGGCTACTCCTGCTGCTGGTTGTGGGCTCCTGGCT
ACTCGCCCGCATCCTGGCTTGACCTATGCCTTCTATAACAAGTCCCGCCGGCTCCAGTGTTC
CACAGCCCCAAAACGGAAGTGGTTTTGGGGTCACCTGGGCTGATCACTCCTACAGAGGAGGGC
TTGAAGGACTCGACCCAGATGTCGGCCACCTATCCCAGGGCTTTACGGTATGGCTGGGTCCCAT
CATCCCTTCATCGTTTTATGCCACCCTGACACCATCCGGTCTATCACCATGCCTCAGTGCCA
TTGCACCAAGGATAATCTCTTCATCAGGTTCTCTGAAGCCCTGGCTGGGAGAAGGGATACTGCTG
AGTGGCGGTGACAAGTGGAGCCGCCACCGTCGGATGCTGACGCCCGCCTTCCATTTCAACATCCT
GAAGTCTATATAACGATCTTCAACAAGAGTGCAAACATCATGCTTGACAAGTGGCAGCACCTGG
CCTCAGAGGGCAGCAGTCGTCTGGACATGTTTGAGCACATCAGCCTCATGACCTTGGACAGTCTA
CAGAAATGCATCTTCAGCTTTGACAGCCATTGTCAGGAGAGGGCCAGTGAATATATTGCCACCAT
CTTGAGGCTCAGTGCCCTTGTAGAGAAAAGAAGCCAGCATATCCTCCAGCACATGGACTTTCTGT
ATTACCTCTCCCATGACGGGCGCGCTTCCACAGGGCCTGCCGCTGGTGCATGACTTCACAGAC
GCTGTCATCCGGGAGCGGCGTCGCACCCTCCCCACTCAGGGTATTGATGATTTTTTCAAAGACAA
AGCCAAGTCCAAGACTTTGGATTTTCATTGATGTGCTTCTGCTGAGCAAGGATGAAGATGGGAAGG
CATTGTCAGATGAGGATATAAGAGCAGAGGCTGACACCTTCATGTTTGGAGGCCATGACACCAG
GCCAGTGGCCTCTCCTGGGTCTGTACAACCTTGCGAGGCACCCAGAATACCAGGAGCGCTGCCG
ACAGGAGGTGCAAGAGCTTCTGAAGGACCGGATCCTAAAGAGATTGAATGGGACGACCTGGCCC
AGCTGCCCTTCCCTGACCATGTGCGTGAAGGAGAGCCTGAGGTTACATCCCCAGCTCCCTTCATC
TCCCGATGCTGCACCCAGGACATTGTTCTCCAGATGGCCGAGTCATCCCCAAAGGCATTACCTG
CCTCATCGATATTATAGGGGTCCATCACAACCCAACTGTGTGGCCGGATCCTGAGGTCTACGACC
CCTTCCGCTTTGACCCAGAGAACAGCAAGGGGAGGTACCTCTGGCTTTTATTCTTTCTCCGCA
GGGCCCAGGAAGTGCATCGGGCAGGCGTTCGCCATGGCGGAGATGAAAGTGGTCTGGCGTTGAT
GCTGCTGCACTTCCGGTTCCTGCCAGACCACACTGAGCCCCGAGGAAGCTGGAATTGATCATGC
GCGCCGAGGGCGGGCTTTGGCTGCGGGTGGAGCCCCTGAATGTAGGCTTGCACTGACTTTCTGAC
CCATCCACCTGTTTTTTTGCAGATTGTCATGAATAAACGGTGTGTCAA

FIGURE 54

MSLLSLPWLGLRPVAMSPWLLLLLVVGSWLLARILAWTYAFYNNCRRLQCFPPKRNWFWGHLG
LITPTEEGLKDSTQMSATYSQGFTVWLGPIIPFIVLCHPDTIRSITNASAAIAPKDNLFIRFLKP
WLGEGILLSGGDKWSRHRMLTPAFHFNILKSYITIFNKSANIMLDKWQHLASEGSSRLDMFEHI
SLMTLDSLQKCIFSFDSHCQERPSEYIATILELSALVEKRSQHILQHMDFLYYLSHDGRRFHRAC
RLVHDFTDAVIRERRRTLPTQGIDDDFFKDKAKSKTLD FIDVLLLSKDEDGKALSDEDIRAEADTF
MFGGHDTTASGLSWVLYNLARHPEYQERCQEVQELLKDRDPKEIEWDDLAQLPFLTMCVKESLR
LHPPAPFISRCCTQDIVLPDGRVIPKGITCLIDIIGVHHNPTVWPDPEVYDPFRFDPENSKGRSP
LAFIPFSAGPRNCIGQAFAMAEMKVVLALMLLHFRFLPDHTEPRRKLELIMRAEGLWLRVEPLN
VGLQ

Important features:

Transmembrane domains:

amino acids 13-32 (type II), 77-102

Cytochrome P450 cysteine heme-iron ligand signature.

amino acids 461-471

N-glycosylation sites.

amino acids 112-116, 168-172

FIGURE 55

ATCGCATCAATTGGGAGTACCATCTTCCTCATGGGACCAGTGAAACAGCTGAAGCGAATGTTTGA
GCCTACTCGTTTGATTGCAACTATCATGGTGCTGTTGTGTTTTGCACTTACCCTGTGTTCTGCCT
TTTGGTGGCATAACAAGGGACTTGCACTTATCTTCTGCATTTTGCACTTTTGGCATTGACGTGG
TACAGCCTTTCCCTTCATACCATTTGCAAGGGATGCTGTGAAGAAGTGTTTGCCGTGTGTCTTGC
ATAATCATGGCCAGTTTTATGAAGCTTTGGAAGGCACATGGACAGAAGCTGGTGGACAGTTTT
GTAACATCTTCGAAACCTCTGTCTTACAGACATGTGCCTTTTATCTTGCAGCAATGTGTGCTT
GTGATTGGAACATTTGAGGGTTACTTTTGGGAAGCAACAATACATTCTCGAACCTGAATGTCAGTA
GCACAGGATGAGAAGTGGGTTCTGTATCTTGTGGAGTGAATCTTCCTCATGTACCTGTTTCCTC
TCTGGATGTTGTCCCACTGAATTCCCATGAATACAAACCTATTCAGCAACAGCAAAAAAAAAAAAA
AA

FIGURE 56

MGPVKQLKRMFEPTRLIATIMVLLCFALTLCSAFWWHNKGLALIFCILQSLALTWYSLSFIPFAR
DAVKKCFVCLA

Important features:

Signal peptide:

amino acids 1-33

Type II fibronectin collagen-binding domain protein.

amino acids 30-72

10063549

FIGURE 57

CGGCTCGAGCTCGAGCCGAATCGGCTCGAGGGGCGAGTGGAGCACCCAGCAGGCCGCCAATGCTCTGTCTGTGCCTG
TACGTGCCGGTCATCGGGGAAGCCAGACCGAGTTCCAGTACTTTAGTTCGAAGGGGCTCCCTGCCGAGCTGAAGTCC
ATTTTCAAGCTCAGTGTCTTTCATCCCTCCAGGAATTCACCTACCGCCAGTGGAGCAGAAAAATTGTACAAGCT
GGAGATAAGGACCTTGATGGGCAGCTAGACTTTGAAGAATTTGTCCATATCTCCAAGATCATGAGAAGAGCTGAGG
CTGGTGTAAAGATTTTGGACAAAAAAGATGATGGACGCATTGACGCGCAGGAGATCATGCAGTCCCTGCCGGACTTG
GGAGTCAAGATATCTGAACAGCAGGCAGAAAAAATTCAGAGCATGGATAAAAACGGCAGCATGACCATCGACTGG
AACGAGTGGAGAGCTACCACCTCCTCCACCCGTTGGAACATCCCGAGATCATCTCTACTGGAAGCATTCCACG
ATCTTTGATGTGGGTGAGAATCTAACGGTCCCGGATGAGTTCACAGTGGAGGAGAGGCAGACGGGGATGTGGTGGAGA
CACCTGGTGGCAGGAGTGGGGCAGGGGCCGTATCCAGAACCTGCACGGCCCCCTGGACAGGCTCAAGGTGCTCATG
CAGGTCCATGCCTCCCGCAGCAACAACATGGGCATCGTTGGTGGCTTCACTCAGATGATTCGAGAAGGAGGGGCCAGG
TCACTCTGGCGGGCAATGGCATCAACGTCTCAAATTGCCCCGAATCAGCCATCAAATTCATGGCTATGAGCAG
ATCAAGCGCCTTGTGGTAGTGACCAGGAGACTCTGAGGATTACGAGAGGCTTGTGGCAGGGTCTTGGCAGGGGCC
ATCGCCCCAGAGCAGCATCTACCCAATGGAGGTCTGAAGACCCGGATGGCGCTGCGGAAGACAGGCCAGTACTCAGGA
ATGCTGGACTGCGCCAGGAGGATCCTGGCCAGAGAGGGGGTGGCCGCTTCTACAAAGGCTATGTCCCCAATGCTG
GGCATCATCCCTTATGCCGGCATCGACCTTGCACTTACGAGAGCGCTCAAGAATGCTGGCTGCAGCACTATGCACTG
AACAGCGCGGACCCCGGGTGTGGTCTCTGGCCTGTGGCACCATTGTCAGTACCTGTGGCCAGCTGGCCAGCTAC
CCCCGGCCCTAGTCAGGACCCGGATGCAGGCGCAAGCCTCTATTAGGGCGCTCCGGAGGTGACCATGAGCAGCCTC
TTCAAACATATCTGCGGACCGAGGGGGCTTCGGGCTGTACAGGGGGCTGGCCCCAATTCATGAAGGTTCATCCA
GCTGTGAGCATCAGTACGTGGTCTACGAGAACCTGAAGATCACCTGGGCGTGCAGTCCGGGTGACGGGGGAGGGC
CGCCCCGCGAGTGGACTCGCTGATCCTGGGCCGAGCCTGGGGTGTGCAGCCATCTCATTCTGTGAATGTGCCAACACT
AAGCTGTCTCGAGCCAAGCTGTGAAACCCCTAGACGCACCCGAGGGAGGGTGGGGAGAGCTGGCAGGCCAGGGCTT
GTCCTGCTGACCCAGCAGACCTCCTGTTGGTTCAGCGAAGACCACAGGCATTCTTAGGGTCCAGGGTCAGCAGG
CTCCGGGCTCACATGTGTAAAGACAGGACATTTCTGCAGTGCCTGCCAATAGTGAGCTTGGAGCCTGGAGGCCGCT
TAGTCTCTCCATTTACCCCTTGACGCCAGCTGTTGGCCACGGCCCTGCCCTCTGGTCTGCCGTGCATCTCCCTGTGC
CCTCTTGCTGCCTGCCTGTCTGCTGAGGTAAGGTGGGAGGAGGGCTACAGCCACATCCACCCCTCGTCCAATCCC
ATAATCCATGATGAAAGGTGAGGTACGTGGCCTCCAGGCCCTGACTTCCCAACCTACAGCATTGACGCCAACTTGGC
TGTGAAGGAAGAGGAAAGGATCTGGCCTTGTGGTCACTGGCATCTGAGCCCTGCTGATGGCTGGGGCTCTCGGGCATG
CTTGGGAGTGCAGGGGGCTCGGGCTGCCTGGCCTGGCTGCACAGAAGCAAGTGCTGGGGCTCATGGTCTCTGAGCT
GGCCTGGACCTGTGAGGATGGGGCCACCTCAGAACCAACTCACTGTCCCACTGTGGCATGAGGGCAGTGGAGCA
CCATGTTTGAGGGCGAAGGGCAGAGCGTTTGTGTCTGGGGAGGGAAGGAAAGGTGTTGGAGCCTTAATTATGG
ACTGTTGGGAAAAGGTTTTGTCCAGAAGGACAGCCGACAAATGAGCGACTTCTGTGCTTCCAGAGGAAGACGAGG
GAGCAGGAGCTTGGCTGACTGCTCAGAGTCTGTTCTGACGCCCTGGGGGTCTCTGTCCAACCCAGCAGGGGGCAGC
GGGACCAGCCCCACATTCCACTTGTGTCACTGCTTGAACCTATTATTATTGTATTATTGTAACAGAGTTATGTCCT
AACTATTTTTATAGATTTGTTTAATTAATAGCTTGTCAATTTCAAGTTCATTTTTATTTCATATTTATGTTTCATGTT
GATTGTACCTTCCCAAGCCCCCAGTGGGATGGGAGGAGGAGGAAAGGGGGCTTGGGCCGCTGCAGTCACATCT
GTCCAGAGAAATTCCTTTGGGACTGGAGGCAGAAAAGCGGCCAGAAAGGCAGCAGCCCTGGCTCCTTTCTTTGGCAG
GTTGGGGAAGGGCTTGGCCCCAGCCTTAGGATTTAGGGTTTGAAGTGGGGGCTGGAGAGAGGGAGGAAACCTCAAT
AACCTTGAAGGTGGAATCCAGTTATTTCTGCGCTGCGAGGGTTCTTTATTTCACTCTTTCTGAATGTCAAGGCAG
TGAGGTGCTCTCACTGTGAATTTGTGGTGGGCGGGGGCTGGAGGAGAGGGTGGGGGGCTGGCTCCCTCCCTCCAGC
CTTCTGCTGCCCTTGTCTTAACAATGCCGGCCAACTGGCGACCTCACGGTTGCACTTCCATTCCACCAGAAATGACCTGA
TGAGGAAATCTTCAATAGGATGCAAGATCAATGCAAAAATTTGTTATATATGAACATATAACTGGAGTCGTCAAAAAG
CAAAATTAAGAAAGAAATGGACGTTAGAAGTTGTCAATTAAGCAGCCTTCTAATAAAGTTGTTTCAAGCTGAAAAAA
AA

FIGURE 58

MLCLCLYVPVIGEATQTEFQYFESKGLPAELKSIFKLSVFIPSQEFSTYRQWKQKIVQAGDKDLDG
QLDFEEFVHYLQDHEKKRLRVFKILDKKNDGRIDAQEI MQSLRDLGVKISEQQAEKILKSMDKNG
TMTIDWNEWRDYHLLHPVENIPEIILYWKHSTIFDVGENLTPDEFTVEERQTGMWWRHLVAGGG
AGAVSRTCTAPLDRLKVLMOVHASRSNNMGIVGGFTQMIREGGARSLWRGNGINVLKIAPESAIK
FMAYEQIKRLVGSDQETLRIHERLVAGSLAGATAQSSIYPMEVLKTRMALRKTGQYSGMLDCARR
ILAREGVAAFYKGYVPNMLGIIPYAGIDLAVYETLKNLQHYAVNSADPGVFVLLACGTMSSTC
GQLASYPLALVRTRMQAQASIEGAPEVTMSSLFKHILRTEGAFGLYRGLAPNFMKVIPAVSISYV
VYENLKITLGVQSR

Important features:

Signal peptide:

amino acids 1-16

Putative transmembrane domains:

amino acids 284-304, 339-360, 376-394

Mitochondrial energy transfer proteins signature.

amino acids 206-215, 300-309

N-glycosylation sites.

amino acids 129-133, 169-173

Elongation Factor-hand calcium-binding protein.

amino acids 54-73, 85-104, 121-140

FIGURE 59

GGAAGGCAGCGGCAGCTCCACTCAGCCAGTACCCAGATACGCTGGGAACCTTCCCCAGCCATGGC
TTCCCTGGGGCAGATCCTCTTCTGGAGCATAATTAGCATCATCATTATTCTGGCTGGAGCAATTG
CACTCATCATTGGCTTTGGTATTTTCAGGGAGACACTCCATCACAGTCACTACTGTCGCCTCAGCT
GGGAACATTGGGGAGGATGGAATCCTGAGCTGCACTTTTGAACCTGACATCAAACCTTCTGATAT
CGTGATACAATGGCTGAAGGAAGGTGTTTTAGGCTTGGTCCATGAGTTCAAAGAAGGCAAAGATG
AGCTGTCGGAGCAGGATGAAATGTTTCAGAGGCCGGACAGCAGTGTTTGCTGATCAAGTGATAGTT
GGCAATGCCTCTTTGCGGCTGAAAAACGTGCAACTCACAGATGCTGGCACCTACAAATGTTATAT
CATCACTTCTAAAGCAAGGGGAATGCTAACCTTGAGTATAAACTGGAGCCTTCAGCATGCCGG
AAGTGAATGTGGACTATAATGCCAGCTCAGAGACCTTGCGGTGTGAGGCTCCCCGATGGTTCCCC
CAGCCCACAGTGGTCTGGGCATCCCAAGTTGACCAGGGAGCCAACTTCTCGGAAGTCTCCAATAC
CAGCTTTGAGCTGAACTCTGAGAATGTGACCATGAAGGTTGTGTCTGTGCTCTACAATGTTACGA
TCAACAACACATACTCCTGTATGATTGAAAATGACATTGCCAAAGCAACAGGGGATATCAAAGTG
ACAGAATCGGAGATCAAAGGCGGAGTCACCTACAGCTGCTAAACTCAAAGGCTTCTCTGTGTGT
CTCTTCTTTCTTTGCCATCAGCTGGGCACTTCTGCCTCTCAGCCCTTACCTGATGCTAAAATAAT
GTGCCTTGGCCACAAAAAAGCATGCAAAGTCATTGTTACAACAGGGATCTACAGAATATTTTAC
CACCAGATATGACCTAGTTTTATATTTCTGGGAGGAAATGAATTCATATCTAGAAGTCTGGAGTG
AGCAAACAAGAGCAAGAAACAAAAAGAAGCCAAAAGCAGAAGGCTCCAATATGAACAAGATAAAT
CTATCTTCAAAGACATATTAGAAGTTGGGAAAATAATTGATGTGAACTAGACAAGTGTGTTAAGA
GTGATAAGTAAATGCACGTGGAGACAAGTGCATCCCCAGATCTCAGGGACCTCCCCCTGCCTGT
CACCTGGGGAGTGAGAGGACAGGATAGTGCATGTTCTTTGTCTCTGAATTTTATGTTATATGTGC
TGTAATGTTGCTCTGAGGAAGCCCCTGGAAAGTCTATCCCAACATATCCACATCTTATATTCCAC
AAATTAAGCTGTAGTATGTACCCTAAGACGCTGCTAATTGACTGCCACTTCGCAACTCAGGGGCG
GCTGCATTTTAGTAATGGGTCAAATGATTCACTTTTTATGATGCTTCCAAAGGTGCCTTGCTTC
TCTTCCCAACTGACAAATGCCAAAGTTGAGAAAAATGATCATAATTTTAGCATAAACAGAGCAGT
CGGGGACACCGATTTTATAAATAAACTGAGCACCTTCTTTTAAACAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 6o

MASLGQILFWSIISIIIIILAGAIALIIGFGISGRHSITVTTVASAGNIGEDGILSCTFEPDIKLS
DIVIQWLKEGVLGLVHEFKEGKDELSEQDEMFRGRTAVFADQVIVGNASRLKNVQLTDAGTYKC
YIITSKKGKNANLEYKTGAFSMPEVNVNDYNASSETLRCEAPRWFPQPTVVWASQVDQGANFSEVS
NTSFELNSENVTMKVSVLYNVTINNTYSCMIENDIAKATGDIKVTESI KRRSHLQLLNSKASL
CVSSFFAISWALLPLSPYLMLK

Important features:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 258-281

N-glycosylation sites.

amino acids 112-116, 160-164, 190-194, 196-200, 205-209, 216-220,
220-224

N-myristoylation sites.

amino acids 52-58, 126-132, 188-194

FIGURE 61

TGACGTCAGAATCACCATGGCCAGCTATCCTTACCGGCAGGGCTGCCCAGGAGCTGCAGGACAAG
CACCAGGAGCCCCCTCCGGGTAGCTACTACCCCTGGACCCCCCAATAGTGGAGGGCAGTATGGTAGT
GGGCTACCCCCTGGTGGTGGTTATGGGGGTCTGCCCCTGGAGGGCCTTATGGACCACCAGCTGG
TGGAGGGCCCTATGGACACCCCAATCCTGGGATGTTCCCTCTGGAAGTCCAGGAGGACCATATG
GCGGTGCAGCTCCCGGGGGCCCCCTATGGTCAGCCACCTCCAAGTTCCTACGGTGCCCAGCAGCCT
GGGCTTTATGGACAGGGTGGCGCCCCCTCCCAATGTGGATCCTGAGGCCTACTCCTGGTTCCAGTC
GGTGGACTCAGATCACAGTGGCTATATCTCCATGAAGGAGCTAAAGCAGGCCCTGGTCAACTGCA
ATTGGTCTTCATTCAATGATGAGACCTGCCTCATGATGATAAACATGTTTGACAAGACCAAGTCA
GGCCGCATCGATGCTACGGCTTCTCAGCCCTGTGGAAATTCATCCAGCAGTGAAGAACCTCTT
CCAGCAGTATGACCGGGACCGCTCGGGCTCCATTAGCTACACAGAGCTGCAGCAAGCTCTGTCCC
AAATGGGCTACAACCTGAGCCCCCAGTTACCCAGCTTCTGGTCTCCCGCTACTGCCCACGCTCT
GCCAATCCTGCCATGCAGCTTGACCGCTTCATCCAGGTGTGCACCCAGCTGCAGGTGTGACAGA
GGCCTTCCGGGAGAAGGACACAGCTGTACAAGGCAACATCCGGCTCAGCTTCGAGGACTTCGTCA
CCATGACAGCTTCTCGGATGCTATTGACCCAACCATCTGTGGAGAGTGGAGTGCACCAGGGACCTT
TCCTGGCTTCTTAGAGTGAGAGAAGTATGTGGACATCTCTCTTTTCTGTCCCTCTAGAAGAAC
ATTCTCCCTTGCTTGATGCAACACTGTTCCAAAGAGGGTGGAGAGTCTGCATCATAGCCACCA
AATAGTGAGGACCGGGGCTGAGGCCACACAGATAGGGGCCTGATGGAGGAGAGGATAGAAGTTGA
ATGTCCTGATGGCCATGAGCAGTTGAGTGGCACAGCCTGGCACCAGGAGCAGGTCTTGTAAATGG
AGTTAGTGTCAGTCAGCTGAGCTCCACCCTGATGCCAGTGGTGAGTGTTCATCGGCCTGTTACC
GTTAGTACCTGTGTTCCCTCACCAGGCCATCCTGTCAAACGAGCCCATTTTCTCAAAGTGAAT
CTGACCAAGCATGAGAGAGATCTGTCTATGGGACCAGTGGCTTGGATTCTGCCACACCATAAAT
CCTTGTGTGTTAACTTCTAGCTGCCTGGGGCTGGCCCTGCTCAGACAAATCTGCTCCCTGGGCAT
CTTTGGCCAGGCTTCTGCCCCCTGCAGCTGGGACCCCTCACTTGCCCTGCCATGCTCTGCTCGGCT
TCAGTCTCCAGGAGACAGTGGTCACCTCTCCCTGCCAATACTTTTTTAATTGCATTTTTTTC
ATTTGGGGCCAAAAGTCCAGTGAAATTGTAAGCTTCAATAAAAGGATGAAACTCTGA

FIGURE 62

MASYPYRQGCPGAAGQAPGAPPGSYYPGPPNSGGQYGSGLPPGGGYGGPAPGGPYGPPAGGGPYG
HPNPGMFPSGTPGGPYGGAAPGGPYGQPPSSYGAQQPGLYGQGGAPPNVDP EAYSWFQSVSDSH
SGYISMKELKQALVNCNWSSFNDETCLMMINMFDKTKSGRIDVYGFSALWKFIQQWKNLFQQYDR
DRSGSISYTELQQALSQMGYNLSPQFTQLLVSRYCPRSANPAMQLDRFIQVCTQLQVLTEAFREK
DTAVQGNIRLSFEDFVTMTASRML

Important features of the protein:

Signal peptide:

amino acids 1-19

N-glycosylation site.

amino acids 147-150

Casein kinase II phosphorylation sites.

amino acids 135-138, 150-153, 202-205, 271-274

N-myristoylation sites.

amino acids 9-14, 15-20, 19-24, 33-38, 34-39, 39-44, 43-48, 61-
66, 70-75, 78-83, 83-88, 87-92, 110-115

FIGURE 63

CAGGATGCAGGGCCGCGTGGCAGGGAGCTGCGCTCCTCTGGGCCTGCTCCTGGTCTGTCTTCATC
TCCCAGGCCTCTTTGCCCGGAGCATCGGTGTTGTGGAGGAGAAAGTTTCCCAAACTTCGGGACC
AACTTGCCCTCAGCTCGGACAACCTTCCTCCACTGGCCCCCTCTAACTCTGAACATCCGCAGCCCGC
TCTGGACCCTAGGTCTAATGACTTGGCAAGGGTTCTCTGAAGCTCAGCGTGCCTCCATCAGATG
GCTTCCACCTGCAGGAGGTTCTGCAGTGCAGAGGTGGCCTCCATCGTGGGGGCTGCCTGCCATG
GATTCTTGCCCCCTGAGGATCCTTGGCAGATGATGGCTGCTGCGGCTGAGGACCGCTGGGGGA
AGCGCTGCCTGAAGAACTCTCTTACCTCTCCAGTGTGCGGCCCTCGCTCCGGGCAGTGGCCCTT
TGCCCTGGGGAGTCTTCTCCCGATGCCACAGGCCTCTCACCTGAGGCTTCACTCCTCCACCAGGAC
TCGGAGTCCAGACGACTGCCCCGTCTAATTCACTGGGAGCCGGGGAAAAATCCTTTCCAACG
CCCTCCCTGGTCTCTCATCCACAGGGTTCTGCCTGATCACCCCTGGGGTACCCTGAATCCCAGTG
TGCTCTGGGGAGGTGGAGGCCCTGGGACTGGTTGGGGAACGAGGCCCATGCCACACCTGAGGGA
ATCTGGGGTATCAATAATCAACCCCCAGGTACCAGCTGGGGAAATATTAATCGGTATCCAGGAGG
CAGCTGGGGAAATATTAATCGGTATCCAGGAGGCAGCTGGGGGAATATTAATCGGTATCCAGGAG
GCAGCTGGGGGAATATTCATCTATACCCAGGTATCAATAACCCATTTCTCCTGGAGTTCTCCGC
CCTCCTGGCTCTTCTTGAACATCCCAGCTGGCTTCCCTAATCCTCCAAGCCCTAGGTTGCAGTG
GGGC TAGAGCACGATAGAGGAAACCCAACATTGGGAGTTAGAGTCCTGCTCCCGCCCTTGCTG
TGTGGGCTCAATCCAGGCCCTGTTAACATGTTCCAGCACTATCCCCACTTTTCAGTGCCTCCCC
TGCTCATCTCCAATAAAATAAAAGCACTTATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAA

FIGURE 64

MQGRVAGSCAPLGLLLVCLHLPGLFARSIGVVEEKVSQNFGTNLPQLGQPSSTGPSNSEHPQPAL
DPRSNDLARVPLKLSVPPSDGFPPAGGSQVQRWPPSWGLPAMDSWPPEDPWQMAAAAEDRLGEA
LPEELSYLSSAAALAPGSGPLPGESSPDATGLSPEASLLHQDSESRRLPRSNSLGAGGKILSQRP
PWSLIHRVLPDHPWGTLNPSVSWGGGGPGTGWGTRMPHPGEGIWGINNQPPGTSGWGNINRYPGGS
WGNINRYPGGSWGNINRYPGGSWGNIHLYPGINNPFPPGVLRPPGSSWNIPAGFPNPPSPRLQWG

Important features of the protein:

Signal peptide:

amino acids 1-26

Casein kinase II phosphorylation sites.

amino acids 56-59, 155-158

N-myristoylation sites.

amino acids 48-53, 220-225, 221-226, 224-229, 247-252, 258-263,
259-264, 269-274, 270-275, 280-285, 281-286, 305-310

[illegible][illegible]

FIGURE 66

MGSGPLVLVLLLTLLGSSHGTPGPMTLQLKLKESFLTNSSYESSFLELLEKLCLLLHLPSTSVTL
HHARSQHHVVCNT

Important features:

Signal peptide:

amino acids 1-19

N-glycosylation site.

amino acids 37-41

N-myristoylation sites.

amino acids 15-21, 19-25, 60-66

FIGURE 67

ACGGACCGAGGGTTCGAGGGAGGGACACGGACCAGGAACCTGAGCTAGGTCAAAGACGCCCGGGC
CAGGTGCCCCGTCGCAGGTGCCCCCTGGCCGGAGATGCGGTAGGAGGGGCGAGCGCGAGAAGCCCC
TTCCTCGGCGCTGCCAACCCGCCACCCAGCCCATGGCGAACCCCGGGCTGGGGCTGCTTCTGGCG
CTGGGCCTGCCGTTCTTGCTGGCCCGCTGGGGCCGAGCCTGGGGGCAAATACAGACCACTTCTGC
AAATGAGAATAGCACTGTTTTGCCTTCATCCACCAGCTCCAGCTCCGATGGCAACCTGCGTCCGG
AAGCCATCACTGCTATCATCGTGGTCTTCTCCCTCTTGCTGCCTTGCTCCTGGCTGTGGGGCTG
GCACTGTTGGTGCGGAAGCTTCGGGAGAAGCGGCAGACGGAGGGCACCTACCGGCCCAGTAGCGA
GGAGCAGTTCTCCCATGCAGCCGAGGCCCGGGCCCCCTCAGGACTCCAAGGAGACGGTGCAGGGCT
GCCTGCCCATCTAGGTCCCCCTCTCCTGCATCTGTCTCCCTTCATTGCTGTGTGACCTTGGGGAAA
GGCAGTGCCCTCTCTGGGCAGTCAGATCCACCCAGTGCTTAATAGCAGGGAAGAAGGTACTTCAA
AGACTCTGCCCCCTGAGGTCAAGAGAGGATGGGGCTATTCACTTTTATATATTTATATAAAATTAG
TAGTGAGATGTAAAAAAAAAAAAAAAAAAAA

FIGURE 68

MANPGLGLLLALGLPFLRLRWGRAWGQIQTTSANENSTVLPSTSSSSSDGNLRPEAITAIIVVFS
LLAALLLAVGLALLVRKLRKREKRQTEGTYRPSSEEQFSHAAEARAPQDSKETVQGCLPI

Important features:

Signal peptide:

amino acids 1-19

Transmembrane domain:

amino acids 56-80

N-glycosylation site.

amino acids 36-40

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 86-90

Tyrosine kinase phosphorylation site.

amino acids 86-94

N-myristoylation sites.

amino acids 7-13, 26-32

FIGURE 69

GCCAGGAATAACTAGAGAGGAACAATGGGGTTATTGAGAGGTTTTGTTTTCTCTTAGTTCTGTGCCTGCTGCCACGAG
TCAAATACTTCCCTCATTAGCTGAATAATAATGGCTTTGAAGATATTGTCATTGTTATAGATCCTAGTGTGCCAGAA
GATGAAAAAATAATTGAACAAATAGAGGATATGGTGACTACAGCTTCTACGTACCTGTTTGAAGCCACAGAAAAAGA
TTTTTTTTTCAAAATGTATCTATATTAATTCCTGAGAAATTGGAAGGAAATCCTCAGTACAAAGGCCAAACATGAA
AACCATAAACATGCTGATGTTATAGTTGCACACCTACACTCCCAGGTAGAGATGAACCATACACCAAGCAGTTCCACA
GAATGTGGAGAGAAAGGCGAATACATTCACTTCACCCTGACCTTCTACTTGGAAAAAACAAATGAATATGGACCA
CCAGGCAAACTGTTTGTCCATGAGTGGGCTCACCTCCGGTGGGGAGTGTGATGAGTACAATGAAGATCAGCCTTTC
TACCGTGCTAAGTCAAAAAAATCGAAGCAACAAGGTGTTCCGCAGGTATCTCTGGTAGAAATAGAGTTTATAAGTGT
CAAGGAGGCGCTGTCTTAGTAGAGCATCGAATTGATTCTACAACAAACTGTATGGAAAAGATTGTCAATTTCTTT
CCTGATAAAGTACAACAGAAAAAGCATCCATAATGTTATGCAAAGTATTGATTCTGTTGTTGAATTTTGTAAACGAA
AAAACCCATAATCAAGAAGCTCCAAGCCTACAAAACATAAAGTGCAATTTTAGAAGTACATGGGAGGTGATTAGCAAT
TCTGAGGATTTTAAAAACACCATACCCATGGTGACACCACCTCCTCCACCTGTCTTCTCATTGCTGAAGATCAGTCAA
AGAATTGTGTGCTTAGTTCTTGATAAGTCTGGAAGCATGGGGGGTAAGGACCGCTAAATCGAATGAATCAAGCAGCA
AAACATTTCTGCTGCAGACTGTTGAAAATGGATCCTGGGTGGGGATGGTTCACTTTGATAGTACTGCCACTATTGTA
AATAAGCTAATCCAAATAAAAGCAGTGATGAAAGAAACACACTCATGGCAGGATTACCTACATATCCTCTGGGAGGA
ACTTCCATCTGCTCTGGAATTAATATGCATTTTCAGGTGATTGGAGAGCTACATTTCCAACTCGATGGATCCGAAGTA
CTGCTGCTGACTGATGGGGAGGATAACACTGCAAGTTCTTGATTGATGAAGTGAACAAAGTGGGGCCATTGTTTCAT
TTTATTGCTTTGGGAAGAGCTGCTGATGAAGCAGTAATAGAGATGAGCAAGATAACAGGAGGAAAGTCAATTTTATGTT
TCAGATGAAGCTCAGAACATGGCCTCATTGATGCTTTTGGGGCTCTTACATCAGGAAATACTGATCTCTCCAGAAG
TCCCTTCAGCTCGAAAGTAAGGATTAACACTGAATAGTAATGCCTGGATGAACGACACTGTCATAATTGATAGTACA
GTGGGAAAGGACACGTTCTTTCTCATCACATGGAACAGTCTGCTCCCAAGTATTTCTCTGCGGATCCCAAGTGAACA
ATAATGGAATTTTACAGTGGATGCAACTTCCAAATGGCCTATCTCAGTATTCAGGAACTGCAAGGTGGGCACT
TGGGCATACAATCTTCAAGCCAAAGCGAACCAGAAACATTAACTATTACAGTAACCTTCTCGAGCAGCAAAATCTTCT
GTGCTCCAATCACAGTGAATGCTAAATGAATAAGGACGTAAACAGTTTCCCAGCCCAATGATTGTTTACGCAGAA
ATTCTACAAGGATATGTACCTGTTCTTGGAGCCAATGTGACTGCTTTCATTGAATCACAGAATGGACATACAGAAGTT
TTGGAATTTTGGATAATGGTGCAGGCGCTGATTCTTTCAAGAATGATGGAGTCTACTCCAGGTATTTTACAGCATAT
ACAGAAAATGGCAGATATAGCTTAAAGTTTCGGGCTCATGGAGGAGCAAACTGCCAGGCTAAAATACGGCCTCCA
CTGAATGAGCCGCTACATACCAGGCTGGGTAGTGAACGGGGAATTAAGCAAAACCCGCAAGACCTGAAATTTGAT
GAGGATACTCAGACCACCTTGGAGGATTTTCCGCGAACAGCATCCGGAGGTGCATTTGTGGTATCACAGTCCCAAGC
CTTCCCTTGCCTGACCAATACCCACCAAGTCAAATCACAGACCTTGATGCCACAGTTCATGAGGATAAGATTATTCTT
ACATGGACAGCACCAGGAGATAATTTGATGTTGAAAAAGTTCAACGTTATATCATAAGAATAAGTGAAGTATTCTT
GATCTAAGAGACAGTTTTGATGATGCTCTCAAGTAATACTACTGATCTGTCAACAAAGGAGGCCAACTCCAAGGAA
AGCTTTGCATTTAAACCAGAAAATATCTCAGAAGAAAATGCAACCCACATATTATTGCCATTAAAAGTATAGATAAA
AGCAATTTGACATCAAAAGTATCCAACATTGCACAAGTAACCTTTGTTTATCCCTCAAGCAATCCTGATGACATTGAT
CCTACACCTACTCCTACTCCTACTCCTACTGATAAAAGTCATAATTCTGGAGTTAATTTTCTACGCTGGTATTG
TCTGTGATTGGGCTGTTGTAATTGTTAACTTTATTTTAAAGTACCACCATTTGAACCTTAACGAAGAAAAAATCTTC
AAGTAGACCTAGAAGAGAGTTTTTAAAAACAAAACAAATGTAAGTAAAGGATATTTCTGAATCTTAAATTCATCCCAT
GTGTGATCATAACTCATAAAAATAATTTAAGATGTCGGAAGAGGATACTTTGATTAAATAAAACACTCATGGATA
TGTAATAACTGTCAAGATTAAAAATTAATAGTTTCATTTATTTGTTATTTTATTGTAAGAAATAGTGATGAACAAAG
ATCCTTTTTTCTACTGATACCTGGTTGTATATTATTTGATGCAACAGTTTTCTGAAATGATATTTCAAATTCATCAA
GAAATTAATCATCTATCTGAGTAGTCAAAATACAAGTAAAGGAGAGCAAAATAACAACATTTGGAAGAAAAA
AA

FIGURE 70

MGLFRGFVFLVLCLLHQSNSTSFIKLNNNGFEDIVIVIDPSVPEDEKIEQIEDMVTASTYLFE
ATEKRFFFKNVSILIPENWKENPQYKRPKHENHKHADVIVAPPTLPGRDEPYTKQFTECGEKGEY
IHFTPDLLLGGKKQNEYGPPGKLFVHEWAHLRWGVFDEYNEDQPFYRAKSKKIEATRCISAGISGRN
RVYKCQGGSCLSRACRIDSTTKLYGKDCQFFPDQVTEKASIMFMQSIDSVVEFCNEKTHNQEAP
SLQNIKCNFRSTWEVISNSEDFKNTIPMVTPPPPPVFSLKISQRIVCLVLDKSGSMGGKDRLNR
MNQAAKHFLQLTVENGSWVGMVHFDSTATIVNKLIQIKSSDERNTLMAGLPTYPLGGTSICSGIK
YAFQVIGELHSQLDGSEVLLLLTDGEDNTASSCIDEVKQSGAIVHFIALGRAADEAVIEMSKITGG
SHFYVSDEAQNNGLIDAFGALTSGNTDLSQKSLQLESKGLTLNSNAWMNDTVIIDSTVGKDTFFL
ITWNSLPSPISLWDPSTIMENFTVDATSKMAYLSIPGTAKVGTWAYNLQAKANPETLTITVTSR
AANSSVPPITVNAKMNKDVNSFSPMIVYAEILQGYVPVLGANVTAFIESQNGHTEVLELLDNGA
GADSEKNDGVYSRYFTAYTENGRYSLKVRAGGANTARLKLRPPLNRAAYIPGWVNGEIEANPP
RPEIDEDTQTTLDEFSRTASGGAFVVSQVPSLPLPDQYPPSQITDLDATVHEDKIIILTWTAPGDN
FDVGKVQRYIIIRISASILDRDSFDDALQVNTDLSPEANSKESFAFKPENISEENATHIFIAI
KSIDKSNLTSKVSANIAQVTLFIPQANPDDIDPTPTPTPTPTPKSHNSGVNISTLVLSVIGSVVI
VNFILSTTI

Signal peptide:

amino acids 1-21

Putative transmembrane domains:

amino acids 284-300, 617-633

Leucine zipper pattern.

amino acids 469-491, 476-498

N-glycosylation site.

amino acids 20-24, 75-79, 340-344, 504-508, 542-546, 588-592,
628-632, 811-815, 832-836, 837-841, 852-856, 896-900

FIGURE 71

CTCCTTAGGTGGAAACCTGGGAGTAGAGTACTGACAGCAAAGACCGGGAAGACCATACGTCCCCGGGCGAGGGGTGA
CAACAGGTGTCATCTTTTGTATCTCGTGTGTGGCTGCCTTCTTATTTCAAGGAAAGACGCCAAGGTAATTTTGACCCA
GAGGAGCAATGATGTAGCCACCTCCTAACCTTCCCTTCTTGAACCCCAAGTTATGCCAGGATTTACTAGAGAGTGTCA
ACTCAACCAGCAAGCGGCTCCTTCGGCTTAACCTGTGGTTGGAGGAGAGAACCTTTGTGGGGCTGCGTTCTCTTAGCA
GTGCTCAGAAGTGACTTGCCTGAGGGTGGACCAGAAGAAAGGTAAGGTCCTTCTGTGCTTGGCTGCACATCAGGAA
GGCTGTGATGGGAATGAAGGTGAAAACCTGGAGATTTCATTCAGTCATTGCTTCTGCCTGCAAGATCATCTTTAAA
AGTAGAGAAGCTGCTCTGTGTGGTGGTTAACTCCAAGAGGCAGAACTCGTTCTAGAAGGAAATGGATGCAAGCAGCTC
CGGGGGCCCCAAACGCATGCTTCCTGTGGTCTAGCCCAAGGGAAGCCCTTCCGTGGGGGGCCCCGGCTTTGAGGGATGCC
ACCGGTTCTGGACGCATGGCTGATTCTGAATGATGATGTTTCGCCGGGGGCTGCTTGCCTGGATTTCGCCGGTGGTG
GTTTTGCTGGTGTCTCTGTGTGCTATCTGTCTGTACATGTTGGCTGCACCCCAAGGAGACGAGGAGCAG
CTGGCACTGCCAGGGCCAAACAGCCCAAGGGAAGGAGGTTACCAGGCCCTCTTTCAGGAGTGGGAGGAGCAGCAC
CGCACTACGTGAGCAGCTGAAGCGGCAGATCGCACAGCTCAAGGAGGAGCTGCAGGAGAGGAGTGAAGCAGCTCAGG
AATGGGCAGTACCAAGCCAGCGATGCTGCTGGCTGGGTCTGGACAGGAGCCCCCAGAGAAAACCCAGGCGCACCTC
CTGGCCTTCTGCACTCGCAGGTGGACAAGGCAGAGTGAATGCTGGCTCAAGCTGGCCACAGAGTATGCAGCAGTG
CCTTTGATAGCTTACTCTACAGAGGTGTACAGCTGGAGACTGGCCTTACCCGCCACCCGAGGAGAACCTGTG
AGGAAGGACAAGCGGATGAGTTGGTGGAGCCATTGAATCAGCCTTGGAGACCTGAACAATCTGCAGAGAACAGC
CCCAATCACCCTCTTACACGGCCTCTGATTTCATAGAAGGGATCTACCGAACAGAAAGGACAAAGGGACATTGTAT
GAGCTCACCTTCAAAGGGGACACAAACACGAATTCAAACGGCTCATCTTATTTGACCATTCAGCCCCATCATGAA
GTGAAAATGAAAAGCTCAACATGGCCAAACAGCTTATCAATGTTATCGTGCCTCTAGCAAAAAGGGTGGACAAGTTC
CGGCAGTTCATGCAGAAATTCAGGGAGATGTGCATTGAGCAGGATGGGAGAGTCCATCTCACTGTTGTTTCTTTGGG
AAAGAAGAAATAAATGAAGTCAAAGGAATCTTGAAACACTTCCAAAGCTGCCAATTGAGGAACCTTACCTTCATC
CAGCTGAATGGAGAAATTTCTCGGGGAAAGGGACTTGATGTTGGAGCCCGCTTCTGGAAGGGAAGCAACGTCCTTCTC
TTTTCTGTGATGTGGACATCTACTTCACATCTGAATTCCTCAATACGTGTAGGCTGAATACACAGCCAGGGAAGAAG
GTATTTTATCCAGTTCTTTTCAGTCAGTACAATCCTGGCATAATATACGGCCACCATGATGCAGTCCCTCCCTTGGAA
CAGCAGCTGGTCATAAAGAAAGAACTGGATTTTGGAGAGACTTTGGATTTGGGATGACGTGTGAGTATCGGTGACAG
TTCATCAATATAGTGGTGGTTGATCTGGACATCAAAGGCTGGGGCGGAGAGGATGTGACCTTTATCGCAAGTATCTC
CACAGCAACCTCATAGTGGTACGGACGCTGTGGCAGGACTCTTCCACCTCTGGCATGAGAAGCGCTGCATGGACGAG
CTGACCCCGAGCAGTACAAGATGTGCATGCAGTCCAAGGCCATGAACGAGGCATCCACGGCCAGCTGGGCATGCTG
GTGTTGAGGCACGAGATAGAGGCTCACCTTCGCAACAGAAACAGAAAGACAAGTAGCAAAAAACATGAAGTCCCAGA
GAAGGATTTGGGAGACACTTTTCTTTTCTTTTGCATTAAGTGGCTGCAACAGAGAAAAGACTTCCATAAAA
GGACGACAAAAGAAATGGACTGATGGGTGAGAGATGAGAAAGCCTCCGATTTCTCTGTGGGCTTTTACACAGA
AATCAAAATCTCCGCTTTGCTGCAAAAGTAACCCAGTTGACCCCTGTGAAGTGTCTGCAAAAGGCAGAAATGCTTGTG
AGATTATAAGCCTAATGGTGTGGAGGTTTGTGTTTGTGTTTACAACTGAGACCTGTTGTTTGTGTGCTCATTTGA
AATATTCATGATTTAAGAGCAGTTTGTAAAAAATTCATTAGCATGAAAGGCAAGCATATTTCTCTCATATGAATGA
GCCTATCAGCAGGGCTCTAGTTTCTAGGAATGCTAAAATATCAGAAGGCAGGAGAGGATAGGCTTATATGATACT
AGTGAGTACATTAAGTAAATAAAATGGACCAGAAAAGAAAAGAACCATAAATATCGTGTATATTTTCCCCAAGAT
TAACCAAAAATAATCTGCTTATCTTTTGGTTGCTTTTAACTGTCTCCGTTTTTTCTTTTATTTAAAAATGCACT
TTTTTTCCCTTGTGAGTTATAGTCTGCTTATTTAATTACCACTTTGCAAGCCTTACAAGAGAGCACAAGTTGGCCCTAC
ATTTTATATTTTAAAGAGATACTTTGAGATGCATTATGAGAACTTTCAGTTCAAAGCATCAAATGATGCCATAT
CCAAGGACATGCCAAATGCTGATTCTGTGAGGCACTGAATGTGAGGCACTGAGACATAGGGAAGGAATGGTTGTACT
AATACAGACGTACAGATACTTTCTCTGAAGAGTATTTTGAAGAGGAGCAACTGAACACTGGAGGAAAAGAAATGAC
ACTTTCTGCTTTACAGAAAAGGAACTCATTGAGACTGGTGATATCGTGATGTACCTAAAAGTCAGAAACCAATTTT
CTCCTCAGAAGTAGGACCGCTTTCTTACCTGTTTAAATAAACCAAGTATACCGTGTGAACCAACATCTCTTTTC
AAAACAGGGTGTCTCTCTGCTTCTGGCTTCCATAAGAAGAAATGGAGAAAATATATATATATATATATATATATTGT
GAAAGATCAATCCATCTGCCAGAACTAGTGGGATGGAAGTTTTGCTACATGTTATCCACCCAGGCCAGGTGGGAAG
TAACTGAATATTTTTTAAATTAAGCAGTTCTACTCAATCACCAGATGCTTCTGAAAATTGCATTTTATTACCATTT
CAAACATTTTTTAAAAATAAATACAGTTAACATAGAGTGGTTTTCTTATTATGTGAAAATTATTAGCCAGCACCAG
ATGCATGAGCTAATATCTCTTTGAGTCTTGTCTGTTTGTCTCACAGTAACTCATTGTTTAAAGCTTCAAGAAC
ATTCAAGCTGTTGGTGTGTTAAAAATGCATTGTATTGATTGTACTGGTAGTTTATGAAATTTAATTAACACAGG
CCATGAATGAAGGTGGTATTGCACAGCTAATAAATATGATTTGTGATATGAA

FIGURE 72

MMVRRGLLAWISRVVLLVLLCCAISVLYMLACTPKGDEEQLALPRANSPTGKEGYQAVLQEW
EQHRNYVSSLKRQIAQLKEELQERSEQLRNGQYQASDAAGLGLDRSPPEKTQADLLAFLHSQV
AEVNAGVKLATEYAAVPFDSFTLQKVYQLETGLTRHPEEKPVKDKRDELVEAIESALET
ENSPNHRPYTASDFIEGIYRTERDKGTLYELTFKGDHKHEFKRLILFRPFSPIMKVNEK
TLINVIVPLAKRVDKFRQFMQNFREMCIEQDGRVHLTVVYFGKEEINEVKGILENTSKA
NFRNFTFIQLNGEFSRGKGLDVGARFWKGSNVLLFFCDVDIYFTSEFLNTCRLNTQPGK
KVFPVLFVSQYNPGIYGHHDVPPLEQQLVIKKETGFWRDFGFGMTCQYRSDFINIGGF
DLDIKGWGGEDVHLYRKYLHNLIVVRTPVRLFHLWHEKRCMDLTPEQYKMCMQSKAMNE
ASHGQLGMLVFRHEIEAHLRKQKQKTSSKKT

Important features:

Signal peptide:

amino acids 1-27

N-glycosylation sites.

amino acids 315-319, 324-328

N-myristoylation sites.

amino acids 96-102, 136-142, 212-218, 311-317, 339-345, 393-399

Amidation site.

amino acids 377-381

[illegible][illegible]

Year	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100
1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100	

FIGURE 75

TGCCGCTGCCGCCGCTGCTGCTGTTGCTCCTGGCGGCGCCTTGGGGACGGGCAGTTCCTGTGTC
TCTGGTGGTTTGCCTAAACCTGCAAACATCACCTTCTTATCCATCAACATGAAGAATGTCCTACA
ATGGACTCCACCAGAGGGTCTTCAAGGAGTTAAAGTTACTTACACTGTGCAGTATTTATCACAA
ATTGGCCCACCAGAGGTGGCACTGACTACAGATGAGAAGTCCATTTCTGTTGCTCCTGACAGCTCC
AGAGAAGTGAAGAGAAATCCAGAAGACCTTCTGTTTCCATGCAACAAATATACTCCAATCTGA
AGTATAACGTGTCTGTGTTGAATACTAAATCAAACAGAACGTGGTCCCAGTGTGTGACCAACCAC
ACGCTGGTGTCTACCTGGCTGGAGCCGAACACTCTTTACTGCGTACACGTGGAGTCCTTCGTCCC
AGGGCCCCCTCGCCGTGCTCAGCCTTCTGAGAAGCAGTGTGCCAGGACTTTGAAAGATCAATCAT
CAGAGTTCAAGGCTAAAATCATCTTCTGGTATGTTTTGCCCATATCTATTACCGTGTTTCTTTTT
TCTGTGATGGGCTATTCCATCTACCGATATATCCACGTTGGCAAAGAGAAACACCCAGCAAATTT
GATTTTGATTTATGGAAATGAATTTGACAAAAGATTCTTTGTGCCTGCTGAAAAATCGTGATTA
ACTTTATCACCTCAATATCTCGGATGATTCTAAAATTTCTCATCAGGATATGAGTTTACTGGGA
AAAAGCAGTGATGTATCCAGCCTTAATGATCCTCAGCCCAGCGGGAACCTGAGGCCCCCTCAGGA
GGAAGAGGAGGTGAAACATTTAGGGTATGCTTCGCATTTGATGGAAATTTTTTGTGACTCTGAAG
AAAACACGGAAGGTACTTCTCTCACCCAGCAAGAGTCCCTCAGCAGAACATACCCCCGGATAAA
ACAGTCATTGAATATGAATATGATGTCAGAACCACTGACATTTGTGCGGGCCTGAAGAGCAGGA
GCTCAGTTTGAGGAGGAGGTGTCCACACAAGGAACATTATTGGAGTCGCAGGCAGCGTTGGCAG
TCTTGGGCCCCGAAACGTTACAGTACTCATAACCCCTCAGCTCCAAGACTTAGACCCCTGGCG
CAGGAGCACACAGACTCGGAGGAGGGGCCGGAGGAAGAGCCATCGACGACCTGGTCTGACTGGGA
TCCCCAAACTGGCAGGCTGTGTATTCCTTCGCTGTCCAGCTTCGACCAGGATTCAGAGGGCTGCG
AGCCTTCTGAGGGGGATGGGCTCGGAGAGGAGGGTCTTCTATCTAGACTCTATGAGGAGCCGGCT
CCAGACAGGCCACCAGGAGAAAATGAAACCTATCTCATGCAATTCATGGAGGAATGGGGTTATA
TGTGCAGATGGAAAACTGATGCCAACACTTCCTTTTGCTTTTGTTCCTGTGCAAACAAGTGAG
TCACCCCTTTGATCCCAGCCATAAAGTACCTGGGATGAAAGAAGTTTTTCCAGTTTGTCTAGTGT
CTGTGAGAATTACTTATTTCTTTTCTATTCTCATAGCACGTGTGTGATTGGTTCATGCATGTA
GGTCTCTTAACAATGATGGTGGGCCTCTGGAGTCCAGGGGCTGGCCGGTTGTTCTATGCAGAGAA
AGCAGTCAATAAATGTTTGCCAGACTGGGTGCAGAATTTATTCAGGTGGGTGT

FIGURE 76

MSYNGLHQRVFKELKLLTLCSSQIGPPEVALTTDEKSISSVLTAPEKWKRNPEDLPVSMQQIY
SNLKYNVSVLNTKSNRTWSQCVTNHTLVLTWLEPNTLYCVHVESFVPGPPRAQPSEKQCARTLK
DQSSEFKAKIIFWYVLPISITVFLFSVMGYSIYRYIHVGKEKHPANLILYGNEDKRFFVPAEK
IVINFITLNISSDDSKISHQDMSLLGKSSDVSSLNDPQPSGNLRPPQEEEEVKHLGYASHLMEIFC
DSEENTEGTSLTQQESLSRTIPDPKTVIEYEDVRTTDICAGPEEQELSLQEEVSTQGTLLSQ
ALAVLGPQTLQYSYTPQLQDLPLAQEHTDSEEGPEEEPSTTLVDWDPQTGRLCIPSLSSFDQDS
EGCEPSEGDGLGEEGLLSRLYEPPAPDRPPGENETYLMQFMEEWGLYVQMEN

Important features:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 140-163

N-glycosylation sites.

amino acids 71-74, 80-83, 89-92, 204-207, 423-426

FIGURE 77

GAGGAGCGGGCCGAGGACTCCAGCGTGCCAGGTCTGGCATCCTGCACTTGCTGCCCTCTGACAC
CTGGGAAGATGGCCGGCCCGTGGACCTTCACCCTTCTCTGTGGTTTGCTGGCAGCCACCTTGATC
CAAGCCACCCTCAGTCCCACTGCAGTTCTCATCCTCGGCCAAAAGTCATCAAAGAAAAGCTGAC
ACAGGAGCTGAAGGACCACAACGCCACCAGCATCCTGCAGCAGCTGCCGCTGCTCAGTGCCATGC
GGGAAAAGCCAGCCGGAGGCATCCCTGTGCTGGGCAGCCTGGTGAACACCGTCCTGAAGCACATC
ATCTGGCTGAAGGTCATCACAGCTAACATCCTCCAGCTGCAGGTGAAGCCCTCGGCCAATGACCA
GGAGCTGCTAGTCAAGATCCCCCTGGACATGGTGGCTGGATTCAACACGCCCTGGTCAAGACCA
TCGTGGAGTTCCACATGACGACTGAGGCCAAAGCCACCATCCGCATGGACACCAGTGCAAGTGGC
CCCACCGCCTGGTCCTCAGTGACTGTGCCACCAGCCATGGGAGCCTGCGCATCCAATGCTGTA
TAAGCTCTCCTTCTGGTGAACGCCTTAGCTAAGCAGGTGATGAACCTCCTAGTGCCATCCCTGC
CCAATCTAGTGAAAAACCAGCTGTGTCCCGTGATCGAGGCTTCCTTCAATGGCATGTATGCAGAC
CTCCTGCAGCTGGTGAAGGTGCCCATTTCCCTCAGCATTGACCGTCTGGAGTTTGACCTTCTGTA
TCCTGCCATCAAGGGTGACACCATTAGCTCTACCTGGGGGCCAAGTTGTTGGACTCACAGGGAA
AGGTGACCAAGTGGTTCAATAACTCTGCAGCTTCCCTGACAATGCCACCCTGGACAACATCCCG
TTCAGCCTCATCGTGAGTCAGGACGTGGTGAAGCTGCAGTGGCTGCTGTGCTCTCTCCAGAAGA
ATTCATGGTCCTGTTGGACTCTGTGCTTCTGAGAGTGCCCATCGGCTGAAGTCAAGCATCGGGC
TGATCAATGAAAAGGCTGCAGATAAGCTGGGATCTACCCAGATCGTGAAGATCCTAACTCAGGAC
ACTCCCGAGTTTTTTATAGACCAAGGCCATGCCAAGGTGGCCCAACTGATCGTGCTGGAAGTGTT
TCCCTCCAGTGAAGCCCTCCGCCCTTTGTTACCCCTGGGCATCGAAGCCAGCTCGGAAGCTCAGT
TTTACACCAAAGGTGACCAACTTATACTCAACTGAATAACATCAGCTCTGATCGGATCCAGCTG
ATGAACTCTGGGATTGGCTGGTTCCAACCTGATGTTCTGAAAAACATCATCACTGAGATCATCCA
CTCCATCCTGCTGCCGAACCAGAATGGCAAATTAAGATCTGGGGTCCCAGTGTCATTGGTGAAGG
CCTTGGGATTCGAGGACGTGAGTCCCTCACTGACCAAGGATGCCCTTGTGCTTACTCCAGCCTCC
TTGTGGAAACCCAGCTCTCCTGTCTCCAGTGAAGACTTGGATGGCAGCCATCAGGGAAGGCTGG
GTCCCAGCTGGGAGTATGGGTGTGAGCTCTATAGACCATCCCTCTCTGCAATCAATAAACAATTG
CCTGTGAAAAA

FIGURE 78

MAGPWTFTLLCGLLAATLIQATLSPTAVLILGPKVKEKLTQELKDNATSILQQLPLLSAMREK
PAGGIPVLGSLVNTVLKHIIWLKVITANILQLQVKPSANDQELLVKIPLDMVAGFNTPLVKTIVE
FHMTEAQATIRMDTSASGPTRLVLSDCATSHGSLRIQLLYKLSFLVNALAKQVMNLLVPSLPNL
VKNQLCPVIEASFNGMYADLLQLVKVPISLSIDRLEFDLLYPAIKGDITQLYLGAALLDSQGKVT
KWFNNSAASLTMPITLDNIPFSLIVSQDVVKAABAVALSPEEFMVLLDSVLPESAHLKSSIGLIN
EKAADKLGSTQIVKILTQDTPEFFIDQGHAKVAQLIVLEVFPSSALRPLFTLGIEASSEAQFYT
KGDQLILNINNISDRIQLMNSGIGWFQPDVLKNIITEIIHSILLPNQNGKLRSGVPVSLVKALG
FEAAESSLTkdALVLTPASLWKPSSPVSQ

Important features of the protein:

Signal peptide:

amino acids 1-21

N-glycosylation sites.

amino acids 48-51, 264-267, 401-404

Glycosaminoglycan attachment site.

amino acids 412-415

LBP / BPI / CETP family proteins.

amino acids 407-457

FIGURE 79

GAGAGAAGTCAGCCTGGCAGAGAGACTCTGAAATGAGGGATTAGAGGTGTTCAAGGAGCAAGAGC
TTCAGCCTGAAGACAAGGGAGCAGTCCCTGAAGACGCTTCTACTGAGAGGTCTGCCATGGCCTCT
CTTGGCCTCCAACCTGTGGGCTACATCCTAGGCCTTCTGGGGCTTTTGGGCACACTGGTTGCCAT
GCTGCTCCCCAGCTGGAAAACAAGTTCTTATGTCGGTGCCAGCATTGTGACAGCAGTTGGCTTCT
CCAAGGCCTCTGGATGGAATGTGCCACACACAGCACAGGCATCACCCAGTGTGACATCTATAGC
ACCCCTTCTGGGCCTGCCCGCTGACATCCAGGCTGCCAGGCCATGATGGTGACATCCAGTGCAAT
CTCCTCCCTGGCCTGCATTATCTCTGTGGTGGGCATGAGATGCACAGTCTTCTGCCAGGAATCCC
GAGCCAAAGACAGAGTGGCGGTAGCAGGTGGAGTCTTTTTTCATCCTTGGAGGCCTCCTGGGATTCT
ATTCCTGTTGCCCTGGAATCTTCATGGGATCTACGGGACTTCTACTCACCAGTGGTGCCTGACAG
CATGAAATTTGAGATTGGAGAGGCTCTTTACTTGGGCATTATTTCTTCCCTGTTCTCCCTGATAG
CTGGAATCATCCTCTGCTTTTCCTGCTCATCCCAGAGAAATCGCTCCAACCTACTACGATGCCTAC
CAAGCCCAACCTCTTGCCACAAGGAGCTCTCCAAGGCCTGGTCAACCTCCCAAAGTCAAGAGTGA
GTTCAATTCTACAGCCTGACAGGGTATGTGTGAAAGAACCAGGGGCCAGAGCTGGGGGGTGGCTG
GGTCTGTGAAAAACAGTGGACAGCACCCCGAGGGCCACAGGTGAGGGACACTACCACTGGATCGT
GTCAGAAGGTGCTGCTGAGGATAGACTGACTTTGGCCATTGGATTGAGCAAAGGCAGAAATGGGG
GCTAGTGTAACAGCATGCAGGTTGAATTGCCAAGGATGCTCGCCATGCCAGCCTTTCTGTTTTCC
TCACCTTGCTGCTCCCTGCCCTAAGTCCCCAACCCCTCAACTTGAAACCCCATTCCTTAAGCCA
GGACTCAGAGGATCCCTTTGCCCTCTGGTTTACCTGGGACTCCATCCCCAAACCCACTAATCACA
TCCCACTGACTGACCCTCTGTGATCAAAGACCCCTCTCTCTGGCTGAGGTTGGCTCTTAGCTCATT
GCTGGGGATGGGAAGGAGAAGCAGTGGCTTTTGTGGGCATTGCTCTAACCTACTTCTCAAGCTTC
CCTCCAAAGAACTGATTGGCCCTGGAACCTCCATCCCACTCTTGTTATGACTCCACAGTGTCCA
GACTAATTTGTGCATGAACTGAAATAAAACCATCCTACGGTATCCAGGGAACAGAAAGCAGGATG
CAGGATGGGAGGACAGGAAGGCAGCCTGGGACATTTAAAAAATA

FIGURE 8o

MASLGLQLVGYYLGLLGLLGLTLVAMLLPSWKTSSYVGASIVTAVGFSKGLWMECATHSTGITQCD
IYSTLLGLPADIQAAQAMMTSSAISSLACIISVVGMRCTVFCQESRAKDRVAVAGGVFFILGGL
LGFIPVAVNNLHGILRDFYSPLVPDSMKFEIGEALYLGIISSLFSLIAGIILCFSCSSQRNRSNYY
DAYQAQPLATRSSPRPGQPPKVKSEFNSYSLTGYV

Important features of the protein:

Signal peptide:

amino acids 1-24

Transmembrane domains:

amino acids 82-102, 117-140, 163-182

N-glycosylation site.

amino acids 190-193

PMP-22 / EMP / MP20 family proteins.

amino acids 46-59

FIGURE 81

CCCACGCGTCCGCGCCTCTCCCTTCTGCTGGACCTTCCTTCGTCTCTCCATCTCTCCCTCCTTTC
CCCAGGTTCTCTTTCCACCTTTCTTCTTCCACCTTAGACCTCCCTTCCTGCCCTCCTTTCTCCT
GCCCACGCTGCTTCCTGGCCCTTCTCCGACCCGCTCTAGCAGCAGACCTCCTGGGGTCTGTGG
GTTGATCTGTGGCCCTGTGCCCTCCGTGTCTTTTCGTCTCCCTTCCTCCCGACTCCGCTCCCGG
ACCAGCGGCCTGACCCTGGGGAAGGATGGTTCCTCCGAGGTGAGGGTCTCTCCTCCTTGCTGGGA
CTCGCGCTGCTCTGGTTCCCCCTGGACTCCCACGCTCGAGCCCGCCAGACATGTTCTGCCTTTT
CCATGGGAAGAGATACTCCCCGGCGAGAGCTGGCACCCCTACTTGAGCCACAAGGCCTGATGT
ACTGCCTGCGCTGTACCTGCTCAGAGGGCGCCCATGTGAGTTGTTACCGCTCCACTGTCCGCT
GTCCACTGCCCCAGCCTGTGACGGAGCCACAGCAATGCTGTCCAAGTGTGTGGAACCTCACAC
TCCCTCTGGACTCCGGGCCCCACAAAGTCTGCCAGCACACGGGACCATGTACCAACACGGAG
AGATCTTCAGTGCCCATGAGCTGTTCCCTCCCGCCTGCCCAACCAGTGTGTCTCTGCAGCTGC
ACAGAGGGCCAGATCTACTGCGGCCTCACAACCTGCCCGAACCAGGTGCCAGCACCCCTCCC
ACTGCCAGACTCTGCTGCCAAGCCTGCAAAGATGAGGCAAGTGAAGCAATCGGATGAAGAGGACA
GTGTGCAGTCGCTCCATGGGGTGAGACATCCTCAGGATCCATGTTCCAGTGATGCTGGGAGAAAG
AGAGGCCCCGGGCACCCAGCCCCACTGGCCTCAGCGCCCTCTGAGCTTCATCCCTCGCCACTT
CAGACCAAGGGAGCAGGCAGCACAACCTGTCAAGATCGTCTGAAGGAGAAACATAAGAAAGCCT
GTGTGCATGGCGGGAAGACGTACTCCACGGGGAGGTGTGGCACCCGGCCTTCCGTGCCTTCGGC
CCCTTGCCCTGCATCCTATGCACCTGTGAGGATGGCCGCCAGGACTGCCAGCGTGTGACCTGTCC
CACCGAGTACCCCTGCCGTACCCCGAGAAAGTGGCTGGGAAGTGTGCAAGATTTGCCAGAGG
ACAAAGCAGACCCTGCCACAGTGAGATCAGTTCTACCAGGTGTCCAAGGCACCGGGCCGGGTC
CTCGTCCACACATCGGTATCCCCAAGCCAGACAACCTGCGTCGCTTTGCCCTGGAACACGAGGC
CTCGGACTTGGTGGAGATCTACCTCTGGAAGCTGGTAAAAGATGAGGAACTGAGGCTCAGAGAG
GTGAAGTACCTGGCCCAAGGCCACACAGCCAGAATCTTCCACTTGAATCAGATCAAGAAAGTCAG
GAAGCAAGACTTCCAGAAAGAGGCACAGCACTTCCGACTGCTCGCTGGCCCCCAGGAGGTCCT
GGAACGTCTTCCTAGCCAGACCCCTGGAGCTGAAGGTCACGGCCAGTCCAGACAAAGTGACCAAG
ACATAACAAAGACCTTAACAGTTGCAGATATGAGCTGTATAATTGTTGTTATTATATATTAATAAA
TAAGAAGTTGCATTACCCTCAAAAAAAAAAAAAAAAAAAAAA

FIGURE 82

MVPEVRVLSSLLGLALLWFPPLDSHARARPDMFCLFHGKRYSPGESWHPYLEPQGLMYCLRCTCSE
GAHVSCYRLHCPPVHCPQPVTEPQQCCPKCVPHTPSGLRAPPKSCQHNGTMYQHGEI FSAHELF
PSRLPNQCVLCSCTEGQIYCGLTTCPEPGCPAPLPLPDSCCQACKDEASEQSDEEDSVQSLHGVR
HPQDPCSSDAGRKRGPPTAPTGLSAPLSFIPRHFRPKGAGSTTVKIVLKEKHKKACVHGKTY
HGEVWHPAFRAFGPLPCILCTCEDGRQDCQRVTCPTTEYPCRHPKVAGKCKICPEDKADPGHSE
ISSTRCPKAPGRVLVHTSVSPSPDNLRFALEHEASDLVEIYLWKLVKDEETEAQRGEVPGPRPH
SQNLPLDSDQESQEARLPERGTALPTARWP PRRSLERLPSDPGAEGHGQSRQSDQDITKT

Signal peptide:

amino acids 1-25

FIGURE 83

GACAGCTGTGTCTCGATGGAGTAGACTCTCAGAACAGCGCAGTTTGCCCTCCGCTCACGCAGAGCCTCTCC
 GTGGCTTCCGCACCTTGAGCATTAGGCCAGTTCTCCTTCTCTCTAATCCATCCGTCACCTCTCCTGTCA
 TCCGTTTCCATGCCGTGAGGTCCATTACAGAACACATCCATGGCTCTCATGCTCAGTTTGGTTCTGAGTC
 TCCTCAAGCTGGGATCAGGGCAGTGGCAGGTGTTTGGGCCAGACAAGCCTGTCCAGGCCTTGGTGGGGGAG
 GACGCAGCATTCTCCTGTTTCTGTCTCCTAAGACCAATGCAGAGGCCATGGAAGTGCGGTTCTTCAGGGG
 CCAGTTCTCTAGCGTGGTCCACCTCTACAGGGACGGGAAGGACCAGCCATTTATGCAGATGCCACAGTATC
 AAGGCAGGACAAAACCTGGTGAAGGATTCTATTGCGGAGGGGCGCATCTCTCTGAGGCTGGAAAACATTACT
 GTGTTGGATGCTGGCCTCTATGGGTGCAGGATTAGTTCCAGTCTTACTACCAGAAGGCCATCTGGGAGCT
 ACAGGTGTGAGCACTGGGCTCAGTTCTCTCATTTCATCACGGGATATGTTGATAGAGACATCCAGCTAC
 TCTGTGCTCCTCGGGCTGGTTCCCCCGGCCACAGCGAAGTGGAAGGTCCACAAGGACAGGATTGTGCC
 ACAGACTCCAGGACAAACAGAGACATGCATGGCCTGTTTGATGTGGAGATCTCTCTGACCGTCCAAGAGAA
 CGCCGGGAGCATATCCTGTTCCATGCGGCATGCTCATCTGAGCCGAGAGGTGGAATCCAGGGTACAGATAG
 GAGATACCTTTTTTCGAGCCTATATCGTGGCACCTGGCTACCAAAGTACTGGGAATACTCTGCTGTGGCCTA
 TTTTTTGGCATTGTGGACTGAAGATTTTCTTCTCAAATTCAGTGGAATAATCCAGGCGGAAGTGGACTG
 GAGAAGAAAGCACGGACAGGCAGAAATTGAGAGACGCCCGGAAACACCGAGTGGAGGTGACTCTGGATCCAG
 AGACGGCTCACCCGAAGCTCTGCGTTTCTGATCTGAAAACGTAAACCCATAGAAAAGCTCCCCAGGAGGTG
 CCTCACTCTGAGAAGAGATTTACAAGGAAGAGTGTGGTGGCTTCTCAGAGTTTCCAAGCAGGGAACATTA
 CTGGGAGGTGGACGGAGGACACAATAAAGGTGGCGCTGGGAGTGTGCCGGGATGATGTGGACAGGAGGA
 AGGAGTACGTGACTTTGTCTCCCGATCATGGGTACTGGGTCTCAGACTGAATGGAGAACATTTGTATTTT
 ACATTAAATCCCCGTTTTATCAGCGTCTTCCCCAGGACCCACCTACAAAATAGGGGTCTTCTCTGGACTA
 TGAGTGTGGGACCATCTCCTTCTTCAACATAAATGACCAGTCCCTTATTTATACCCTGACATGTGGGTTT
 AAGGCTTATTGAGGCCCTACATTGAGTATCCGTCTATAATGAGCAAAATGGAAGTCCCATAGTCATCTGC
 CCAGTCACCCAGGAATCAGAGAAAGAGGCCTCTTGGCAAAGGGCCTCTGCAATCCCAGAGACAAGCAACAG
 TGAGTCTCTCTACAGGCAACACGCCCTTCTCCCCAGGGGTGAAATGTAGGATGAATCACATCCCACAT
 TCTTCTTTAGGGATATTAAGGTCTCTCTCCAGATCCAAAGTCCCGCAGCAGCCGGCCAAGGTGGCTTCCA
 GATGAAGGGGGACTGGCCTGTCCACATGGGAGTCAGGTGTCTGCTGCCCCTGAGCTGGGAGGGAAGAAGG
 CTGACATTACATTTAGTTTGTCTCTCACTCCATCTGGCTAAGTGATCTTGAATACCACCTCTCAGGTGAAG
 AACCCTCAGGAATCCCATCTCACAGGCTGTGGTGTAGATTAAGTAGACAAGGAATGTGAATAATGCTTAG
 ATCTTATTGATGACAGAGTGATCCTAATGGTTTGTTCATTATATTACACTTTCAGTAAAAAAA

FIGURE 84

MALMLSLVLSLLKLGSGQWQVFGPDKPVQALVGEDAAFCFLSPKTNAEAMEVRFFRGQFSSVVH
LYRDGKDQPFMQMPQYQGR TKLVKDSIAEGRISLRLENITVLDAGLYGCRISSQSYQKAIWELQ
VSALGSVPLISITGYVDRDIQLLCQSSGWFPRTAKWKGPQGDLSTDSRTNRDMHGLFDVEISL
TVQENAGSISCSMRHAHLSREVESRVQIGDTFFEPISWHLATKVLGILCCGLFFGIVGLKIFFSK
FQWKIQAE LDWRRKHGQAE LRDARKHAVEVT LDPETAHPKLCVSDLKTVTHR KAPQEVP HSEKRF
TRKSVVASQS FQAGKHYWEVDGGHNKRWRVGVC RDDVDRRKEYVTLS PDHGYWVLR LNLGEHLYFT
LNPRFISVFPRT PPTKIGVFLDYECGTISFFNINDQSLIYTLTCRFEGLLRPYIEYPSYNEQNGT
PIVICPVTQESEKEASWQRASAI PETSNSSESSQATT PFLPRGEM

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 239-255

FIGURE 85

AACAGACGTTCCCTCGCGGCCCTGGCACCTCTAACCCAGACAATGCTGCTGCTGCTGCTGCCCCCT
GCTCTGGGGGAGGGAGAGGGCGGAAGGACAGACAAGTAACTGCTGACGATGCAGAGTTCCGTGA
CGGTGCAGGAAGGCTGTGTGTCCATGTGCCCTGCTCCTTCTCCTACCCCTCGCATGGCTGGATT
TACCCTGGCCCAGTAGTTTCATGGCTACTGGTTCCGGGAAGGGGCCAATACAGACCAGGATGCTCC
AGTGGCCACAAACAACCCAGCTCGGGCAGTGTGGGAGGAGACTCGGGACCGATTCCACCTCCTTG
GGGACCCACATACCAAGAATTGCACCCTGAGCATCAGAGATGCCAGAAGAAGTGATGCGGGGAGA
TACTTCTTTCGTATGGAGAAAGGAAGTATAAAATGGAATTATAAACATCACCGGCTCTCTGTGAA
TGTGACAGCCTTGACCCACAGGCCCAACATCCTCATCCCAGGCACCTTGAGTCCGGCTGCCCCC
AGAATCTGACCTGCTCTGTGCCCTGGGCCTGTGAGCAGGGGACACCCCTATGATCTCCTGGATA
GGGACCTCCGTGTCCCCCTGGACCCCTCCACCACCCGCTCCTCGGTGCTCACCTCATCCCACA
GCCCCAGGACCATGGCACCAGCCTCACCTGTCAGGTGACCTTCCCTGGGGCCAGCGTGACCACGA
ACAAGACCGTCCATCTCAACGTGTCTACCCGCTCAGAACTTGACCATGACTGTCTTCCAAGGA
GACGGCACAGTATCCACAGTCTTGGGAAATGGCTCATCTGTCACTCCCAGAGGGCCAGTCTCT
GCGCCTGGTCTGTGCAGTTGATGCAGTTGACAGCAATCCCCCTGCCAGGCTGAGCCTGAGCTGGA
GAGGCTGACCTGTGCCCTCACAGCCCTCAAACCCGGGGTGCTGGAGCTGCCTTGGGTGCAC
CTGAGGGATGCAGCTGAATTCACCTGCAGAGCTCAGAACCTCTCGGCTCTCAGCAGGTCTACCT
GAACGTCTCCCTGCAGAGCAAAGCCACATCAGGAGTACTCAGGGGGTGGTCGGGGGAGCTGGAG
CCACAGCCCTGGTCTTCTGTCTTCTGCGTCATCTTCGTTGTAGTGAGGTCTGCAGGAAGAAA
TCGGCAAGGCCAGCAGCGGGCGTGGGAGATACGGGCATAGAGGATGCAAACGCTGTGAGGGGTTT
AGCCTCTCAGGGGCCCTGACTGAACCTTGGGCAGAAGACAGTCCCCCAGACCAGCCTCCCCAG
CTTCTGCCCGCTCCTCAGTGGGGGAAGGAGAGCTCCAGTATGCATCCCTCAGCTTCCAGATGGTG
AAGCCTTGGGACTCGCGGGGACAGGAGGCCACTGACACCGAGTACTCGGAGATCAAGATCCACAG
ATGAGAACTGCAGAGACTCACCTGATTGAGGGATCACAGCCCCTCCAGGCAAGGGAGAAGTCA
GAGGCTGATTCTTGTAGAATTAACAGCCCTCAACGTGATGAGCTATGATAACACTATGAATTATG
TGCAGAGTGAAAAGCACACAGGCTTTAGAGTCAAAGTATCTCAAACCTGAATCCACACTGTGCCC
TCCCTTTTATTTTTTAACTAAAAGACAGACAAATTCCTA

FIGURE 86

MLLLLLPLLWGRERAEGQTSKLLTMQSSVTVQEGLCVHVPCSFSYPSHGWIYPGPVVHGYWFREG
ANTDQDAPVATNNPARAVWEETRDRFHLLGDPHTKNCTLSIRDARRSDAGRYFFRMEKGSIKWNY
KHHRLSVNVTALTHRPNILIPGTLESGCPQNLTCVWPWACEQGTPPMISWIGTSVSPDPSTTRS
SVLTLLIPQPQDHGTSLTQVTFPGASVTTNKTVHLNVSYPQNLMTVTFQGDGTVSTVLGNGSSL
SLPEGQSLRLVCAVDAVDSNPPARLSLSWRGLTLCPSQPSNPGVLELPWVHLRDAAEFTCRAQNP
LGSQQVYLNVSLSKATSGVTQGVGGAGATALVFLSFCVIFVVVRSCRKKSARPAAGVGDGTGIE
DANAVRGASQGPLTEPWAEDSPDPQPPASARSSVGEGLQYASLSFQMVKPWDSRGQEATDTE
YSEIKIHR

Signal peptide:

amino acids 1-15

Transmembrane domain:

amino acids 351-370

FIGURE 87

AGAAAGCTGCACTCTGTTGAGCTCCAGGGCCAGTGGAGGGAGGGAGTGAAGGAGCTCTCTGTAC
CCAAGGAAAGTGCAGCTGAGACTCAGACAAGATTACAATGAACCAACTCAGCTTCCTGCTGTTTC
TCATAGCGACCACCAGAGGATGGAGTACAGATGAGGCTAATACTTACTTCAAGGAATGGACCTGT
TCTTCGTCTCCATCTCTGCCCAGAAGCTGCAAGGAAATCAAAGACGAATGTCTAGTGCATTTGA
TGGCTGTATTTTCTCCGCACTGAGAATGGTGTATCTACCAGACCTTCTGTGACATGACCTCTG
GGGGTGGCGGCTGGACCCTGGTGGCCAGCGTGCATGAGAATGACATGCGTGGGAAGTGCACGGTG
GGCGATCGCTGGTCCAGTCAGCAGGGCAGCAAAGCAGACTACCCAGAGGGGGACGGCAACTGGGC
CAACTACAACACCTTTGGATCTGCAGAGGCGGCCACGAGCGATGACTACAAGAACCCTGGCTACT
ACGACATCCAGGCCAAGGACCTGGGCATCTGGCACGTGCCCAATAAGTCCCCCATGCAGCACTGG
AGAAACAGCTCCCTGCTGAGGTACCGCACGGACACTGGCTTCCTCCAGACACTGGGACATAATCT
GTTTGGCATCTACCAGAAATATCCAGTGAAATATGGAGAAGGAAAGTGTGGACTGACAACGGCC
CGGTGATCCCTGTGGTCTATGATTTTGGCGACGCCAGAAAACAGCATCTTATTACTCACCTAT
GGCCAGCGGGAATTCAGTGCAGGATTTGTTTCAAGTTCAGGGTATTTAATAACGAGAGAGCAGCCAA
CGCCTTGTGTGCTGGAATGAGGGTCACCGGATGTAACACTGAGCATCACTGCATTGGTGGAGGAG
GATACTTTCCAGAGGCCAGTCCCCAGCAGTGTGGAGATTTTCTGGTTTTGATTGGAGTGGATAT
GGAACATCATGTTGGTTACAGCAGCAGCCGTGAGATAACTGAGGCAGCTGTGCTTCTATTCTATCG
TTGAGAGTTTTGTGGGAGGGAACCCAGACCTCTCTCCCAACCATGAGATCCCAAGGATGGAGAA
CAACTTACCCAGTAGCTAGAATGTAAATGGCAGAAGAGAAAACAATAAATCATATTGACTCAAGA
AAAAAA

FIGURE 88

MNQLSFLLFLIATTRGWSTDEANTYFKEWTCSSSPSLPRSCKEIKDECPSAFDGLYFLRTENGVI
YQTFCDMTSGGGGWTLVASVHENDMRGKCTVGDRWSSQQGSKADYPEGDGNWANYNTFGSAEAAT
SDDYKNPGYYDIQAKDLGIWHVPNKSPMQHWRNSSLLRYRTDTGFLQTLGHNLFGIYQKYPVKYG
EGKCWTDNGPVI PVVYDFGDAQKTASYISPYGQREFTAGFVQFRVFNNERAANALCAGMRVTGCN
TEHHCIGGGGYFPEASPPQCGDFSGFDWSGYGTHVGYSSSREITEAAVLLFYR

Important features:

Signal peptide:

amino acids 1-16

N-glycosylation site.

amino acids 163-167

Glycosaminoglycan attachment sites.

amino acids 74-78, 289-293

N-myristoylation sites.

amino acids 76-82, 115-121, 124-130, 253-259, 292-298

FIGURE 89

CTAGATTTGTCGGCTTGCGGGGAGACTTCAGGAGTCGCTGTCTCTGAACTTCCAGCCTCAGAGAC
CGCCGCCCTTGTCCTCCGAGGGCCATGGGCCGGGTCTCAGGGCTTGTGCCCTCTCGCTTCCTGACG
CTCCTGGCGCATCTGGTGGTCGTCATCACCTTATTCTGGTCCCGGGACAGCAACATACAGGCCTG
CCTGCCTCTCACGTTACCCCCGAGGAGTATGACAAGCAGGACATTCAGCTGGTGGCCGCGCTCT
CTGTCACCCTGGGCCTCTTTGCAGTGGAGCTGGCCGGTTTCCTCTCAGGAGTCTCCATGTTCAAC
AGCACCCAGAGCCTCATCTCCATTGGGGCTCACTGTAGTGCATCCGTGGCCCTGTCCTTCTTCAT
ATTCGAGCGTTGGGAGTGCACCTACGTATTGGTACATTTTTGTCTTCTGCAGTGCCCTTCCAGCTG
TCACTGAAATGGCTTTATTTCGTCACCGTCTTTGGGCTGAAAAAGAAACCCTTCTTGATTACCTTCA
TGACGGGAACCTAAGGACGAAGCCTACAGGGGCAAGGGCCGCTTCGTATTCCTGGAAGAAGGAAG
GCATAGGCTTCGGTTTTCCCTCGGAACTGCTTCTGCTGGAGGATATGTGTTGGAATAATTACG
TCTTGAGTCTGGGATTATCCGCATTGTATTTAGTGCTTTGTAATAAAATATGTTTTGTAGTAACA
TTAAGACTTATATACAGTTTTAGGGGACAATTAAAAAAAAAAAA

FIGURE 90

MGRVSGLVPSRFLTLLAHLVVVITLFWSRDSNIQACPLPTFTPEEYDKQDIQLVAALSVTLGLFA
VELAGFLSGVSMFNSTQSLISIGAHCSASVALSFFIFERWECTTYWYIFVFCALPAVTEMALFV
TVFGLKKKPF

Transmembrane domain:

amino acids 12-28 (type II), 51-66, 107-124

FIGURE 91

CTGGGACCCCGAAAAGAGAAGGGGAGAGCGAGGGGACGAGAGCGGAGGAGGAAGATGCAACTGAC
TCGCTGCTGCTTCGTGTTCTTGGTGCAGGGTAGCCTCTATCTGGTCATCTGTGGCCAGGATGATG
GTCCTCCCGGCTCAGAGGACCCTGAGCGTGATGACCACGAGGGCCAGCCCCGGCCCCGGGTGCCT
CGGAAGCGGGGCCACATCTCACCTAAGTCCCCCCCCATGGCCAATTCCACTCTCCTAGGGCTGCT
GGCCCCGCCTGGGGAGGCTTGGGGCATTCTTGGGCAGCCCCCAACGCCCCGAACCACAGCCCCC
CACCCTCAGCCAAGGTGAAGAAAATCTTTGGCTGGGGCGACTTCTACTCCAACATCAAGACGGTG
GCCCTGAACCTGCTCGTCACAGGGAAGATTGTGGACCATGGCAATGGGACCTTCAGCGTCCACTT
CCAACACAATGCCACAGGCCAGGGAAACATCTCCATCAGCCTCGTGCCCCCAGTAAAGCTGTAG
AGTTCCACCAGGAACAGCAGATCTTCATCGAAGCCAAGGCCTCCAAAATCTTCAACTGCCGGATG
GAGTGGGAGAAGGTAGAACGGGGCCGCCGACCTCGCTTTGCACCCACGACCCAGCCAAGATCTG
CTCCCAGACCACGCTCAGAGCTCAGCCACCTGGAGCTGCTCCCAGCCCTTCAAAGTCGTCTGTG
TCTACATCGCCTTCTACAGCACGGACTATCGGCTGGTCCAGAAGGTGTGCCAGATTACAACCTAC
CATAGTGATACCCCCCTACTACCCATCTGGGTGACCCGGGGCAGGCCACAGAGGCCAGGCCAGGGC
TGGAAGGACAGGCCTGCCCCATGCAGGAGACCATCTGGACACCGGGCAGGGAAGGGGTGGGCCTC
AGGCAGGGAGGGGGTGGAGACGAGGAGATGCCAAGTGGGGCCAGGGCCAAGTCTCAAGTGGCAG
AGAAAGGGTCCCAAGTGCTGGTCCCAACCTGAAGCTGTGGAGTGACTAGATCACAGGAGCACTGG
AGGAGGAGTGGGCTCTCTGTGCAGCCTCACAGGGCTTTGCCACGGAGCCACAGAGAGATGCTGGG
TCCCCAGGCCTGTGGGCAGGCCGATCAGTGTGGCCCCAGATCAAGTCATGGGAGGAAGCTAAGC
CCTTGGTTCTTGCCATCCTGAGGAAAGATAGCAACAGGGAGGGGGAGATTTTCATCAGTGTGGACA
GCCTGTCAACTTAGGATGGATGGCTGAGAGGGCTTCCTAGGAGCCAGTCAGCAGGGTGGGGTGGG
GCCAGAGGAGCTCTCCAGCCCTGCCTAGTGGGGCCCTGAGCCCCTTGTCGTGTGCTGAGCATGG
CATGAGGCTGAAGTGGCAACCCTGGGGTCTTTGATGTCTTGACAGATTGACCATCTGTCTCCAGC
CAGGCCACCCCTTTCCAAAATTCCCTCTTCTGCCAGTACTCCCCCTGTACCACCCATTGCTGATG
GCACACCCATCCTTAAGCTAAGACAGGACGATTGTGGTCTCCACACTAAGGCCACAGCCCATC
CGCGTGCTGTGTGTCCCTCTTCCACCCCAACCCCTGCTGGCTCCTCTGGGAGCATCCATGTCCCG
GAGAGGGGTCCCTCAACAGTCAGCCTCACCTGTGACCGGGGTCTCCCGGATCTGGATGGCGC
CGCCCTCTCAGCAGCGGGCACGGGTGGGGCGGGGCCGCGCAGAGCATGTGCTGGATCTGTTC
TGTGTGTCTGTCTGTGGGTGGGGGAGGGGAGGAAGTCTTGTGAAACCGCTGATTGCTGACTTT
TGTGTGAAGAATCGTGTCTTGGAGCAGGAAATAAAGCTTGCCCCGGGGCA

FIGURE 92

MQLTRCCFVFLVQGSLLYLVICGQDDGPPGSEDPERDDHEGQPRPRVPRKRGHISPKSRPMANSTL
LGLLAPPGEAWGILGQPPNRPNHSPPPSAKVKKIFGWGDFYSNIKTVALNLLVTGKIVDHGNGTF
SVHFQHNATGQGNISISLVPPSKAVEFHQEQQIFIEAKASKIFNCRMWEKVERGRRTSLCTHDP
AKICSRDHAQSSATWSCSQPFKVVVCVYIAFYSTDYRLVQKVC PDYNYHSDTPYYPSG

Important features of the protein:

Signal peptide:

amino acids 1-14

N-glycosylation sites.

amino acids 62-65, 127-130, 137-140, 143-146

2-oxo acid dehydrogenases acyltransferase

amino acids 61-71

FIGURE 93

CGGTGGCCATGACTGCGGCCGTGTTCTTCGGCTGCGCCTTCATTGCCTTCGGGCCCTGCGCTCGCC
CTTTATGTCTTCACCATCGCCATCGAGCCGTTGCGTATCATCTTCCTCATCGCCGGAGCTTTCTT
CTGGTTGGTGTCTCTACTGATTTCGTCCCTTGTTTGGTTCATGGCAAGAGTCATTATTGACAACA
AAGATGGACCAACACAGAAATATCTGCTGATCTTTGGAGCGTTTGTCTGTCTATATCCAAGAA
ATGTTCCGATTTGCATATTATAAACTCTTAAAAAAGCCAGTGAAGGTTTGAAGAGTATAAACCC
AGGTGAGACAGCACCCCTCTATGCGACTGCTGGCCTATGTTTCTGGCTTGGGCTTTGGAATCATGA
GTGGAGTATTTTCCTTTGTGAATACCCTATCTGACTCCTTGGGGCCAGGCACAGTGGGCATTCAT
GGAGATTCTCCTCAATTCTTCCTTTATTCAGCTTTCATGACGCTGGTCATTATCTTGCTGCATGT
ATTCTGGGGCATTGTATTTTTTGTGAGTGTGAGAAGAAAAAGTGGGGCATCCTCCTTATCGTTC
TCCTGACCCACCTGCTGGTGTGAGCCAGACCTTCATAAGTTCTTATTATGGAATAAACCTGGCG
TCAGCATTTATAATCCTGGTGCTCATGGGCACCTGGGCATTCTTAGCTGCGGGAGGCAGCTGCCG
AAGCCTGAAACTCTGCCTGCTCTGCCAAGACAAGAAGTTTCTTCTTTACAACCAGCGCTCCAGAT
AACCTCAGGGAACCAGCACTTCCCAAACCGCAGACTACATCTTTAGAGGAAGCACAACTGTGCCT
TTTTCTGAAAATCCCTTTTTCTGGTGGAATTGAGAAAGAAATAAACTATGCAGATA

FIGURE 94

MTAAVFFGCAFIAGFPALALYVFTIAIEPLRIIFLIAGAFFWLVSLLISSLVWFMARVIIDNKDG
PTQKYLLIFGAFVSVYIQEMFRFAYYKLLKKASEGLKSINPGETAPSMRLLAYVSGLGFGIMSGV
FSFVNTLSDSLGPGTVGIHGDSPOFFLYSAEMTLVIILLHVFWGIVFFDGCEKKKWGILLIVLLT
HLLVSAQTFISSYYGINLASAFIILVLMGTWAFLAAGGSCRSLKCLLCQDKNFLLYNQSR

Important features of the protein:

Signal peptide:

amino acids 1-19

Transmembrane domains:

amino acids 32-51, 119-138, 152-169, 216-235

Glycosaminoglycan attachment site.

amino acids 120-123

Sodium:neurotransmitter symporter family protein

amino acids 31-65

FIGURE 95

AATTTTTCACCAGAGTAAACTTGAGAAACCAACTGGACCTTGAGTATTGTACATTTTGCCTCGTG
GACCCAAAGGTAGCAATCTGAAACATGAGGAGTACGATTCTACTGTTTTGTCTTCTAGGATCAAC
TCGGTCATTACCACAGCTCAAACCTGCTTTGGGACTCCCTCCACAAAACCTGGCTCCGGATCAGG
GAACACTACCAAACCAACAGCAGTCAAATCAGGTCTTTCCTTCTTTAAGTCTGATACCATTAACA
CAGATGCTCACACTGGGGCCAGATCTGCATCTGTAAATCCTGCTGCAGGAATGACACCTGGTAC
CCAGACCCACCCATTGACCCTGGGAGGGTTGAATGTACAACAGCAACTGCACCCACATGTGTTAC
CAATTTTGTACACAACCTTGAGCCCAGGGCACTATCCTAAGCTCAGAGGAATTGCCACAAATC
TTCACGAGCCTCATCATCCATTCCCTTGTTCCCGGGAGGCATCCTGCCACCAGTCAGGCAGGGGC
TAATCCAGATGTCCAGGATGGAAGCCTTCCAGCAGGAGGAGCAGGTGTAAATCCTGCCACCCAGG
GAACCCAGCAGGCCGCCTCCCAACTCCCACTGGCACAGATGACGACTTTGCAGTGACCACCCCT
GCAGGCATCCAAAGGAGCACACATGCCATCGAGGAAGCCACCACAGAATCAGCAAATGGAATTCA
GTAAGCTGTTTCAAATTTTTTCAACTAAGCTGCCTCGAATTGGTGATACATGTGAATCTTTATC
ATTGATTATATTATGGAATAGATTGAGACACATTGGATAGTCTTAGAAGAAATTAATTCTTAATT
TACCTGAAAATATTCTTGAAATTCAGAAAATATGTTCTATGTAGAGAATCCCACTTTTAAAAA
CAATAATTCATGGATAAATCTGTCTTTGAAATATAACATTATGCTGCCTGGATGATATGCATAT
TAAACATATTTGGAACCTGGAAA
AAA

FIGURE 96

MRSTILLFCLLGSTRSLPQLKPALGLPPTKLAPDQGTLPNQQSNOVFPSLSLIPLTQM
LTLGPDHLHLLNPAAGMTPGTQTHPLTLGGGLNVQQQLHHPHVLPIFVTQLGAQGTILSSEE
LPQIFTSLSLIIHSLFPGGILPTSQAGANPDVQDGLPAGGAGVNPATQGTPAGRLPTPSG
TDDDFAVTTPAGIQRSTHAIEEATTESANGIQ

Signal peptide:

amino acids 1-16

FIGURE 97

GCTCAAGTGCCCTGCCTTGCCCCACCCAGCCAGCCTGGCCAGAGCCCCCTGGAGAAGGAGCTCT
 CTTCTTGCTTGGCAGCTGGACCAAGGGAGCCAGTCTTGGGCGCTGGAGGGCCTGTCTTGACCATG
 GTCCCTGCCTGGCTGTGGCTGCTTTGTGTCTCCGTCCCCAGGCTCTCCCAAGGCCAGCCTGC
 AGAGCTGTCTGTGAAGTTCCAGAAAATATGGTGGAATTTCCCTTTATACCTGACCAAGTTGC
 CGCTGCCCGTGAGGGGGCTGAAGGCCAGATCGTGCTGTCAAGGGGACTCAGGCAAGGCAACTGAG
 GGCCCATTTGCTATGGATCCAGATTCTGGCTTCTGCTGGTGACCAGGGGCCCTGGACCGAGAGGA
 GCAGGCAGAGTACCAGCTACAGGTACCCTGGAGATGCAGGATGGACATGTCTTGTGGGTCCAC
 AGCCTGTGCTTGTGCACGTGAAGGATGAGAATGACCAGGTGCCCCATTTCTCTCAAGCCATCTAC
 AGAGCTCGGCTGAGCCGGGGTACCAGGCCTGGCATCCCCCTTCTCTTCTTGAGGCTTCAGACCG
 GGATGAGCCAGGCACAGCCAACCTCGGATCTTCGATTCCACATCCTGAGCCAGGCTCCAGGCCAGC
 CTTCCCCAGACATGTTCAGCTGGAGCCTCGGCTGGGGGGCTCTGGCCCTCAGCCCCAAGGGGAGC
 ACCAGCCTTGACCACGCCCTGGAGAGGACCTACCAGCTGTTGGTACAGGTCAAGGACATGGGTGA
 CCAGGCCTCAGGCCACCAGGCCACTGCCACCGTGAAGTCTCCATCATAGAGAGCACCTGGGTGT
 CCCTAGAGCCTATCCACCTGGCAGAGAATCTCAAAGTCTTATACCCGCACCACATGGCCAGGTA
 CACTGGAGTGGGGGTGATGTGCATATCACCTGGAGAGCCATCCCCCGGGACCCCTTTGAAGTGAA
 TGCAGAGGGAAACCTCTACGTGACCAGAGAGCTGGACAGAGAAGCCCAGGCTGAGTACCTGCTCC
 AGGTGCGGGCTCAGAATTCCCATGGCGAGGACTATGCGGGCCCTCTGGAGCTGCACGTGCTGGTG
 ATGGATGAGAATGACAACGTGCCATCTCTGCCCTCCCCGTGACCCACAGTCAAGTCAAGTCAAGT
 CAGTCCACCAGGTACTGAAGTGACTAGACTGTGACAGAGGATGCAGATGCCCCCGGCTCCCCCA
 ATTCCACGTTGTGTATCAGCTCCTGAGCCCTGAGCCTGAGGATGGGGTAGAGGGGAGAGCCTTC
 CAGGTGGACCCCACTTCAGGCAGTGTGACGCTGGGGGTGCTCCCACTCCGAGCAGGCCAGAACAT
 CCTGCTTCTGGTGCTGGCCATGGACCTGGCAGGCGCAGAGGGTGGCTTCAGCAGCACGTGTGAAG
 TCGAAGTCGAGTCACAGATATCAATGATCACGCCCTGAGTTCATCACTTCCAGATTGGGCCT
 ATAAGCCCTCCCTGAGGATGTGGAGCCCGGGACTCTGGTGGCCATGCTAACAGCCATTGATGCTGA
 CCTCGAGCCCGCCTTCCGCCTCATGGATTTTGCCATTGAGAGGGGAGACACAGAAGGGACTTTTG
 GCCTGGATTGGGAGCCAGACTCTGGGCATGTTAGACTCAGACTCTGCAAGAACCTCAGTTATGAG
 GCAGCTCCAAGTCATGAGGTGGTGGTGGTGGTGCAGAGTGTGGCGAAGCTGGTGGGGCCAGGCC
 AGGCCCTGGAGCCACCGCCACGGTGAAGTGTGCTAGTGGAGAGAGTGTGCCACCCCCAAGTTGG
 ACCAGGAGAGCTACGAGGCCAGTGTCCCATCAGTGCCCCAGCCGGCTCTTTCTGCTGACCATC
 CAGCCCTCCGACCCCATCAGCCGAACCCCTCAGGTTCTCCCTAGTCAATGACTCAGAGGGCTGGCT
 CTGCATTGAGAAATTTCCGGGGAGGTGCACACCGCCAGTCCCTGCAGGGCGCCAGCCTGGGG
 ACACCTACACGGTGTCTGTGGAGGCCAGGATACAGCCCTGACTCTTGCCCTGTGCCCTCCCAA
 TACCTCTGCACACCCCGCCAAGACCATGGCTTGATCGTGAGTGGACCCAGCAAGGACCCGATCT
 GGCCAGTGGGCACGGTCCCTACAGCTTACCCCTTGGTCCCAACCCACGGTGAACGGGATTGGC
 GCCTCCAGACTCTCAATGGTTCCCATGCCTACCTACCTTGGCCCTGCATTGGGTGGAGCCACGT
 GAACACATAATCCCCGTGGTGGTCAAGGACCAATGCCAGATGTGGCAGCTCCTGGTTGAGTGAT
 CGTGTGCTGCTGCAACGTGGAGGGGAGTGCATGCGCAAGGTGGGCCGATGAAGGGCATGCCCCA
 CGAAGCTGTGGCAGTGGGCATCCTTGTAGGCACCCCTGGTAGCAATAGGAATCTTCTCATCCTC
 ATTTTCAACCACTGGACCATGTCAAGGAAGAAGGACCCGGATCAACCAGCAGACAGCGTGCCCT
 GAAGGCGACTGTCTGAATGGCCCAGGCAGCTCTAGCTGGGAGCTTGGCCTCTGGCTCCATCTGAG
 TCCCTGGGAGAGAGCCAGCACCCAGATCCAGCAGGGGACAGGACAGAGTAGAAGCCCCCTCCA
 TCTGCCCTGGGGTGGAGGCACCATCACCATCACCAGGCATGTCTGCAGAGCCTGGACACCACTT
 TATGGACTGCCCATGGGAGTGCTCCAAATGTCAGGGTGTGGTCCCAATAATAAAGCCCCAGAGAA
 CTGGGCTGGGCCCTATGGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAG

FIGURE 98

MVPAWLWLLCVSVPQALPKAQPAELSVEVPENYGGNFPLYLTKLPLPREGAEGQIVLSGDSGKAT
EGPFAMDPDSGFLLVTRALDREEQAEYQLQVTLEMQDGHVLWGPQPVLVHVKDENDQVPHFSQAI
YRARLSRGTRPGIPFLFLEASDRDEPGTANSDLRFHILSQAPAQSPDMFQLEPRLGALALSPKG
STSLDHALERTYQLLVQVKMDGDAQSGHQATATVEVSI IESTWVSLEPIHLAENLKVLYPHHMAQ
VHWSGGDVHYHLESHPPGPFVNAEGNLYVTRELDREAQAEYLLQVRAQNSHGEDIAAPLELHVL
VMDENDNVPICPPRDPTVSIPELSPPGTEVTRLAEDADAPGSPNSHVYQLLSPEPEDGVEGRA
FQVDPTSGSVTLGLVPLRAGQNILLVLAMDLAEGGFSSTCEVEVAVTDINDHAPEFITSQIG
PISLPEDVEPGTLVAMLTADADLEPAFRIMDFAIERGDTEGTFGLDWEPPDSGHVRLRLCKNLSY
EAAPSHEVVVVVQSVAKLVGPGPGGATATVTVLVERVMPPPKLDQESYEASVPISAPAGSFLLT
IQSPDPISRTLRFSLVNDSEGWLCIEKFSGEVHTAQSLQGAQPGDTYTVLVEAQDTALTALPVPS
QYLCTPRQDHGLIVSGPSKDPDLASGHGPYSFTLGPNPTVQRDWRLQTLNGSHAYLTALHWVEP
REHIIPVVVSHNAQMWQLLVRVIVCRCNVEGQCMRKVGRMKGMPTKLSAVGILVGTILVAIGIFLI
LIFTHWTMSRKKDPDQPADSVPLKATV

Signal peptide:

amino acids 1-18

Transmembrane domain:

amino acids 762-784

FIGURE 99

GGCTGACCGTGCTACATTGCCTGGAGGAAGCCTAAGGAACCCAGGCATCCAGCTGCCCACGCCTG
 AGTCCAAGATTCTTCCCAGGAACACAAACGTAGGAGACCCACGCTCCTGGAAGCACCAGCCTTTA
 TCTCTTCACCTTCAAGTCCCCTTTCTCAAGAATCCTCTGTCTTTTGCCCTCTAAAGTCTTGGTAC
 ATCTAGGACCCAGGCATCTTGCTTTCCAGCCACAAAGAGACAGATGAAGATGCAGAAAGGAAATG
 TTCTCCTTATGTTTGGTCTACTATTGCATTTAGAAGCTGCAACAAATTCCAATGAGACTAGCACC
 TCTGCCAACACTGGATCCAGTGTGATCTCCAGTGGAGCCAGCACAGCCACCAACTCTGGGTCCAG
 TGTGACCTCCAGTGGGGTCAGCACAGCCACCATCTCAGGGTCCAGCGTGACCTCCAATGGGGTCA
 GCATAGTCACCAACTCTGAGTTCATACAACCTCCAGTGGGATCAGCACAGCCACCAACTCTGAG
 TTCAGCACAGCGTCCAGTGGGATCAGCATAGCCACCAACTCTGAGTCCAGCACAACTCCAGTGG
 GGGCAGCACAGCCACCAACTCTGAGTCCAGCACACCCTCCAGTGGGGCCAGCACAGTCAACCACT
 CTGGGTCCAGTGTGACCTCCAGTGGAGCCAGCACTGCCACCAACTCTGAGTCCAGCACAGTGTCC
 AGTAGGGCCAGCACTGCCACCAACTCTGAGTCTAGCACACTCTCCAGTGGGGCCAGCACAGCCAC
 CAACTCTGACTCCAGCACAACTCCAGTGGGGCTAGCACAGCCACCAACTCTGAGTCCAGCACAA
 CCTCCAGTGGGGCCAGCACAGCCACCAACTCTGAGTCCAGCACAGTGTCCAGTAGGGCCAGCACT
 GCCACCAACTCTGAGTCCAGCACAACTCCAGTGGGGCCAGCACAGCCACCAACTCTGAGTCCAG
 AACGACCTCCAATGGGGCTGGCACAGCCACCAACTCTGAGTCCAGCACGACCTCCAGTGGGGCCA
 GCACAGCCACCAACTCTGACTCCAGCACAGTGTCCAGTGGGGCCAGCACTGCCACCAACTCTGAG
 TCCAGCACGACCTCCAGTGGGGCCAGCACAGCCACCAACTCTGAGTCCAGCACGACCTCCAGTGG
 GGCTAGCACAGCCACCAACTCTGACTCCAGCACAACTCCAGTGGGGCCGGCACAGCCACCAACT
 CTGAGTCCAGCACAGTGTCCAGTGGGATCAGCACAGTCAACCAATTCTGAGTCCAGCACACCCTCC
 AGTGGGGCCAACACAGCCACCAACTCTGAGTCCAGTACGACCTCCAGTGGGGCCAACACAGCCAC
 CAACTCTGAGTCCAGCACAGTGTCCAGTGGGGCCAGCACTGCCACCAACTCTGAGTCCAGCACAA
 CCTCCAGTGGGGTCCAGCACAGCCACCAACTCTGAGTCCAGCACAACTCCAGTGGGGCTAGCACA
 GCCACCAACTCTGACTCCAGCACAACTCCAGTGGGGCCAGCACAGCCACCAACTCTGAGTCTAG
 CACAGTGTCCAGTGGGATCAGCACAGTCAACCAATTCTGAGTCCAGCACAACTCCAGTGGGGCCA
 ACACAGCCACCAACTCTGGGTCCAGTGTGACCTCTGCAGGCTCTGGAACAGCAGCTCTGACTGGA
 ATGCACACAACTTCCCATAGTGCATCTACTGCAGTGAAGTGAAGCAAAGCCTGGTGGGTCCCTGGT
 GCCGTGGGAAATCTTCTCATCACCTGGTCTCGGTGTGGCGGCCGTGGGGCTCTTTGCTGGGC
 TCTTCTTCTGTGTGAGAAACAGCCTGTCCCTGAGAAACACCTTTAACACAGCTGTCTACCACCT
 CATGGCCTCAACCATGGCCTTGGTCCAGGCCCTGGAGGGAATCATGGAGCCCCCACAGGCCAG
 GTGGAGTCCCTAACTGGTTCTGGAGGAGACCAGTATCATCGATAGCCATGGAGATGAGCGGGAGGA
 ACAGCGGGCCCCTGAGCAGCCCCGGAAGCAAGTGCCGCATTCTTCAGGAAGGAAGAGACCTGGGCA
 CCCAAGACCTGGTTTCTTTTTCATTTCATCCCAGGAGACCCCTCCAGCTTTGTTTGAATCCTGAA
 AATCTTGAAGAAGGTATTCCTCACCTTTCTTGCTTTTACCAGACACTGGAAAGAGAATACTATAT
 TGCTCATTTAGCTAAGAAATAAATACATCTCATCTAACACACACGACAAAGAGAAGCTGTGCTTG
 CCCCAGGGTGGGTATCTAGCTCTGAGATGAATCAGTTATAGGAGAAAACCTCCATGCTGGACTC
 CATCTGGCATTCAAAATCTCCACAGTAAATCCAAAGACCTCAAAAAAAAAAAAAAAAAAAAAA
 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 100

MKMQKGNVLLMFGLLLHLEAATNSNETSTSA NTGSSV ISSGASTATNSGSSVTSSGVSTATISGS
SVTSNGVSIVTNSEFHTTSSGISTATNSEFSTASSGISIATNSESSTTSSGASTATNSESSTPSS
GASTVTNSGSSVTSSGASTATNSESSTVSSRASTATNSESSTLSSGASTATNSDSSTTSSGASTA
TNSESSTTSSGASTATNSESSTVSSRASTATNSESSTTSSGASTATNSESRTTSNGAGTATNSES
STTSSGASTATNSDSSTVSSGASTATNSESSTTSSGASTATNSESSTTSSGASTATNSDSSTTSS
GAGTATNSESSTVSSGISTVTNSESSTPSSGANTATNSESSTTSSGANTATNSESSTVSSGASTA
TNSESSTTSSGVSTATNSESSTTSSGASTATNSDSSTTSSEASTATNSESSTVSSGISTVTNSES
STTSSGANTATNSGSSVTSAGSGTAALTGMHTTSHSASTAVSEAKPGGSLVPWEIFLITLVSVVA
AVGLFAGLFFCVRNLSLRNTFNTAVYHPHGLNHGLGPGPGGNHGAPHRPRWSPNWFWRPVSII
AMEMSGRNSGP

Signal peptide:

amino acids 1-20

Transmembrane domain:

amino acids 510-532

FIGURE 101

GGCCGGACGCCTCCGCGTTACGGGATGAATTAACGGCGGGTTCGCGACGGAGGTTGTGACCCCTA
CGGAGCCCCAGCTTGCCACGCACCCCACTCGGCGTCGCGCGGCGTGCCCTGCTTGTCACAGGTG
GGAGGCTGGAACATATCAGGCTGAAAAACAGAGTGGGTACTCTCTTCTGGGAAGCTGGCAACAAAT
GGATGATGTGATATATGCATTCCAGGGGAAGGGAAATTGTGGTGCTTCTGAACCCATGGTCAATT
AACGAGGCAGTTTCTAGCTACTGCACGTACTTCATAAAGCAGGACTCTAAAAGCTTTGGAATCAT
GGTGTGTCATGGAAGGGATTACTTTATACTGACTCTGTTTTGGGGAAGCTTTTTGGAAGCATTT
TCATGCTGAGTCCCTTTTTACCTTTGATGTTTGTAACCCATCTTGGTATCGCTGGATCAACAAC
CGCCTTGTTGGCAACATGGCTCACCTACCTGTGGCATTATTGGAGACCATGTTTGGTGTAAGT
GATTATAACTGGGGATGCATTTGTTCTGGAGAAAGAAGTGTCATTATCATGAACCATCGGACAA
GAATGGACTGGATGTTCTGTGGAATTGCCTGATGCGATATAGCTACCTCAGATTGGAGAAAATT
TGCCTCAAAGCGAGTCTCAAAGGTGTTCTGGATTGTTGGTGGGCCATGCAGGCTGCTGCCTATAT
CTTCATTATAGGAAATGGAAGGATGACAAGAGCCATTTTGAAGACATGATTGATTACTTTTGTG
ATATTCACGAACCACTTCAACTCCTCATATTCCCAGAAGGGACTGATCTCACAGAAAACAGCAAG
TCTCGAAGTAATGCATTTGCTGAAAAAATGGACTTCAGAAATATGAATATGTTTTACATCCAAG
AACTACAGGCTTTACTTTTGTGGTAGACCGTCTAAGAGAAGGTAAGAACCTTGATGCTGTCCATG
ATATCACTGTGGCGTATCCTCACACATTCTCAATCAGAGAAGCACCTCCTCCAAGGAGACTTT
CCCAGGGAAATCCACTTTCACGTCCACCGGTATCCAATAGACACCCTCCCCACATCCAAGGAGGA
CCTTCAACTCTGGTGCCACAAACGGTGGGAAGAGAAAGAAGAGAGGCTGCGTTCCTTCTATCAAG
GGGAGAAGAATTTTTATTTTACCGGACAGAGTGTGATTCCACCTTGCAAGTCTGAACTCAGGGTC
CTTGTGGTCAAATTGCTCTCTATACTGTATTGGACCCTGTTTCAGCCCTGCAATGTGCCTACTCAT
ATATTTGTACAGTCTTGTTAAGTGGTATTTTATAATCACCATTGTAATCTTTGTGCTGCAAGAGA
GAATATTTGGTGGACTGGAGATCATAGAACTTGCAATGTTACCGACTTTTACACAAACAGCCACAT
TTAAATTCAAAGAAAAATGAGTAAGATTATAAGGTTTGCCATGTGAAAACCTAGAGCATATTTTG
GAAATGTTCTAAACCTTTCTAAGCTCAGATGCATTTTTCATGACTATGTCGAATATTTCTTACT
GCCATCATTTATTTGTTAAAGATATTTTGCACCTTAATTTTGTGGGAAAAATATTGCTACAATTTT
TTTAATCTCTGAATGTAATTTTCGATACTGTGTACATAGCAGGGAGTGATCGGGGTGAAATAACTT
GGGCCAGAATATTATTAAACAATCATCAGGCTTTTAAA

FIGURE 102

MHSRGREIVVLLNPWSINEAVSSYCTYFIKQDSKSGIMVSWKGIYFILTLFWGSFFGSI FMLS P
FLPLMFVNPSWYRWINNRLVATWLTLPVALLETMFGVKVIITGDAFVPGERSVIIMNHRTRMDWM
FLWNCLMRYSYLRLEKICLKASLKGVPFGFGWAMQAAAYIFIHRKWKDDKSHFEDMIDYFCDIHEP
LQLLIFPEGTDLTENSKSRSNFAEKNGLQKYEYVLHPRTTGFTFVVDRLREGKNLDAVHDITVA
YPHNIPQSEKHLQGDFFPREIHFHVHRYPIDTLPTSKEDLQLWCHKRWEEKERLRSFYQGEKNF
YFTGQSVIPPCKSELRVLVVKLLSILYWTLFSPAMCLLIYLSLVKWYFIITIVIFVLQERIFGG
LEIIEIACYRLLHKQPHLNSKKNE

Important features of the protein:

Signal peptide:

amino acids 1-22

Transmembrane domains:

amino acids 44-63, 90-108, 354-377

FIGURE 103

CGGCTCGAGCGGCTCGAGTGAAGAGCCTCTCCACGGCTCCTGCGCCTGAGACAGCTGGCCTGACC
TCCAAATCATCCATCCACCCCTGCTGTCTCTGTTTTTCATAGTGTGAGATCAACCCACAGGAATA
TCCATGGCTTTTGTGCTCATTTTGGTTCTCAGTTTCTACGAGCTGGTGTGAGGACAGTGGCAAGT
CACTGGACCGGGCAAGTTTGTCCAGGCCTTGGTGGGGGAGGACGCCGTGTTCTCCTGCTCCCTCT
TTCCTGAGACCAGTGCAGAGGCTATGGAAGTGGGTTCTTCAGGAATCAGTTCCATGCTGTGGTC
CACCTCTACAGAGATGGGGAAGACTGGGAATCTAAGCAGATGCCACAGTATCGAGGGAGAAGTGA
GTTTGTGAAGGACTCCATTGCAGGGGGGCGTGTCTCTTAAGGCTAAAAACATCACTCCCTCGG
ACATCGGCCTGTATGGGTGCTGGTTTCAGTTCCAGATTACGATGAGGAGGCCACCTGGGAGCTG
CGGGTGGCAGCACTGGGCTCACTTCTCTCATTTCCATCGTGGGATATGTTGACGGAGGTATCCA
GTTACTCTGCCTGTCTCAGGCTGGTTCCCCAGCCACAGCCAAGTGGAAAGGTCCACAAGGAC
AGGATTTGTCTTCAGACTCCAGAGCAAATGCAGATGGGTACAGCCTGTATGATGTGGAGATCTCC
ATTATAGTCCAGGAAAATGCTGGGAGCATATTGTGTTCCATCCACCTTGCTGAGCAGAGTCATGA
GGTGAATCCAAGGTATTGATAGGAGAGACGTTTTTCCAGCCCTCACCTTGGCGCCTGGCTTCTA
TTTTACTCGGGTTACTCTGTGGTGCCCTGTGTGGTGTGTCATGGGGATGATAATTGTTTTCTTC
AAATCCAAAGGAAAAATCCAGGCGGAAGTGGACTGGAGAAGAAAGCACGGACAGGCAGAATTGAG
AGACGCCCCGAAACACGCAGTGGAGGTGACTCTGGATCCAGAGACGGCTCACCCGAAGCTCTGCG
TTTCTGATCTGAAACTGTAAACCATAGAAAAGCTCCCCAGGAGGTGCCCTCACTCTGAGAAGAGA
TTTACAAGGAAGAGTGTGGTGGCTTCTCAGGGTTTCCAAGCAGGGAGACATTACTGGGAGGTGGA
CGTGGGACAAAATGTAGGGTGGTATGTGGGAGTGTGTCGGGATGACGTAGACAGGGGGAAGAACA
ATGTGACTTTGTCTCCCAACAATGGGTATTGGGTCTCAGACTGACAACAGAACATTTGTATTTT
ACATTCAATCCCCATTTTATCAGCCTCCCCCCCAGCACCCCTCCTACACGAGTAGGGGTCTTCT
GGACTATGAGGGTGGGACCATCTCCTTCTTCAATACAAATGACCAGTCCCTTATTTATACCCTGC
TGACATGTCAGTTTGAAGGCTTGTGAGACCCTATATCCAGCATGCGATGTATGACGAGGAAAAG
GGGACTCCCATATTATATGTCCAGTGTCTTGGGGATTGAGACAGAGAAGACCCTGCTTAAAGGGC
CCCACACCACAGACCAGACACAGCCAAGGGAGAGTGTCTCCGACAGGTGGCCCCAGCTTCTCT
CCGGAGCCTGCGCACAGAGAGTACGCCCCCTCTCTTTAGGGAGCTGAGGTTCTTCTGCCC
TGAGCCCTGCAGCAGCGGCAGTACAGCTTCCAGATGAGGGGGGATTGGCCTGACCCTGTGGGAG
TCAGAAGCCATGGCTGCCCTGAAGTGGGGACGGAATAGACTCACATTAGGTTTAGTTTGTGAAAA
CTCCATCCAGCTAAGCGATCTTGAACAAGTCACAACCTCCCAGGCTCCTCATTTGCTAGTCACGG
ACAGTGATTCTGCCTCACAGGTGAAGATTAAAGAGACAACGAATGTGAATCATGCTTGCAGGTT
TGAGGGCACAGTGTGTTGCTAATGATGTGTTTTATATTATACATTTCCCACCATAAACTCTGTT
TGCTTATTCCACATTAATTTACTTTTCTCTATACCAAATCACCCATGGAATAGTTATTGAACACC
TGCTTTGTGAGGCTCAAAGAATAAAGAGGAGGTAGGATTTTTCTACTGATTCTATAAGCCCAGCAT
TACCTGATACCAAAACCAGGCAAAGAAAACAGAAGAAGAGGAAGGAAAACCTACAGGTCCATATCC
CTCATTAACACAGACACAAAAATTTAAATAAAATTTTAAACAAATTAACATAAATAATATATTA
AAGATGATATATAACTACTCAGTGTGGTTTGTCCCAAAATGCAGAGTTGGTTTAAATTTAAAT
ATCAACCAGTGTAATTACAGCACATTAATAAAGTAAAAAAGAAAACCATAAAAAAAAAAAAAA

FIGURE 104

MAFVLILVLSFYELVSGQWQVTGPGKFVQALVGEDAVFSCSLFPETSAEAMEVRFFRNQFHAVVH
LYRDGEDWESKQMPQYRGRTEFVKDSIAGGRVSLRLKNITPSDIGLYGCWFSSQIYDEEATWELR
VAALGSLPLISIVGYVDGGIQLLCLSSGWFPQPTAKWKGPQGDLSSDSRANADGYSLYDVEISI
IVQENAGSILCSIHAEQSHEVESKVLIGETFFQSPWRLASILLGLLCGALCGVVMGMIIVFFK
SKGKIQAELDWRRKHGQAEIRDARKHAVEVTLDPETAHPKLCVSDLKTVTHRKAPQEVPHSEKRF
TRKSVVASQGFQAGRHYWEVDVGQNVGWYGVCRDDVDRGKNNVTLSPNNGYWVLRLTTEHLYFT
FNPHFISLPPSTPPTRVGVFLDYEGGTISFFNTNDQSLIYTLLTCQFEGLLRPYIQHAMYDEEKG
TPIFICPVSWG

Signal peptide:

amino acids 1-17

Transmembrane domains:

amino acids 131-150, 235-259

FIGURE 105

CCTTCACAGGACTCTTCATTGCTGGTGGCAATGATGTATCGGCCAGATGTGGTGAGGGCTAGGAAAAGAG
TTTGTGGGAACCTGGGTTATCGGCCTCGTCATCTTCATATCCCTGATTGTCTGGCAGTGTGCATTGGA
CTCACTGTTTCATTATGTGAGATATAATCAAAGAAGACCTACAATTACTATAGCACATTGTCATTTACAAC
TGACAAACTATATGCTGAGTTTGGCAGAGAGGCTTCTAACAATTTTACAGAAATGAGCCAGAGACTTGAAT
CAATGGTGAAAAATGCATTTTATAAATCTCCATTAAGGGAAGAATTTGTCAAGTCTCAGGTTATCAAGTTC
AGTCAACAGAAGCATGGAGTGTGGCTCATATGCTGTTGATTGTAGATTTCACTCTACTGAGGATCCTGA
AACTGTAGATAAAATTGTTCAACTTGTTTTACATGAAAAGCTGCAAGATGCTGTAGGACCCCTAAAGTAG
ATCCTCACTCAGTTAAATTAATAAATCAACAAGACAGAAACAGACAGCTATCTAAACCATTGCTGCGGA
ACACGAAGAAGTAAACTCTAGGTCAGAGTCTCAGGATCGTTGGTGGGACAGAAGTAGAAGAGGGTGAATG
GCCCTGGCAGGCTAGCCTGCAGTGGGATGGGAGTCATCGCTGTGGAGCAACCTTAATTAATGCCACATGGC
TTGTGAGTGTGCTCACTGTTTACAACATATAAGAACCCTGCCAGATGGACTGCTTCCTTTGGAGTAACA
ATAAACCTTCGAAAATGAAACGGGGTCTCCGGAGAATAATTGTCCATGAAAAATACAAACACCCATCACA
TGACTATGATATTTCTCTTGCAGAGCTTTCTAGCCCTGTTCCCTACACAAATGCAGTACATAGAGTTTGTCT
TCCCTGATGCATCCTATGAGTTTCAACCAGGTGATGTGATGTTTGTGACAGGATTTGGAGCACTGAAAAAT
GATGGTTACAGTCAAATCATCTTCGACAAGCACAGGTGACTCTCATAGACGCTACAACCTGCAATGAACC
TCAAGCTTACAATGACGCCATAACTCCTAGAATGTTATGTGCTGGCTCCTTAGAAGGAAAAACAGATGCAT
GCCAGGGTGACTCTGGAGGACCACTGGTTAGTTCAGATGCTAGAGATATCTGGTACCTTGCTGGAATAGTG
AGCTGGGGAGATGAATGTGCGAAACCAACAAGCCTGGTGTATATACTAGAGTTACGGCCTTGCGGGACTG
GATTACTTCAAAAAGTGGTATCTAAAGAGACAAAAGCCTCATGGAACAGATAACATTTTTTTTGTTTTTTG
GGTGTGGAGGCCATTTTTAGAGATACAGAATTGGAGAAGACTTGCAAAACAGCTAGATTTGACTGATCTCA
ATAAAGTGTGCTTGATGCATGTATTTCTTCCAGCTCTGTTCCGCACGTAAGCATCCTGCTTCTGCCA
GATCAACTCTGTCTGTGAGCAATAGTTGAACTTTATGTACATAGAGAAATAGATAATACAATATTAC
ATTACAGCCTGTATTCATTTGTTCTCTAGAAGTTTGTGAGAATTTGACTTGTGACATAAATTTGTAAT
GCATATATACAATTTGAAGCACTCCTTTTCTTCAGTTCCTCAGCTCCTCTCATTTTCAGCAAATATCCATTT
TCAAGGTGCAGAACAGGAGTGAAAGAAAAATATAAGAAGAAAAAATCCCCTACATTTTATTGGCACAGAA
AAGTATTAGGTGTTTTCTTAGTGAATATTAGAAATGATCATATTCATTATGAAAGGTCAAGCAAAGACA
GCAGAATACCAATCACTTCATCATTTAGGAAGTATGGGAACAAAGTTAAGGAAGTCCAGAAAGAAGCCAAG
ATATATCCTTATTTTCATTTCCAAACAACACTACTATGATAAATGTGAAGAAGATTCTGTTTTTTGTGACCT
ATAATAATTATACAACTTCATGCAATGTACTTGTCTAAGCAAATTAAGCAAATATTTATTTAACATTG
TTACTGAGGATGTCAACATATAACAATAAAATATAAATCACCCA

FIGURE 106

MMYRPDVVRARKRVCWEPWVIGLVIFISLIVLAVCIGLTVHYVRYNQKKTYNYYSTLSFTTDKLY
AEFGREASNNFTEMSQRLESMVKNAFYKSPLREEFVKSQVIKFSQQKHGVLAHMLLICRFHSTED
PETVDKIVQLVLHEKLQDAVGPPKVDPHSVKIKKINKTETDSYLNHCCGTRRSKTLGQSLRIVGG
TEVEEGEWPWQASLQWDGSHRCGATLINATWLVSAAHCFTTYKNPARWTASFGVTIKPSKMKRGL
RRIIVHEKYKHPSHDYDISLAELSSPVPTNAVHRVCLPDASYEFQPGDVMFVTGFGALKNDGYS
QNHLRQAQVTLIDATTCNEPQAYNDAITPRMLCAGSLEGKTDACQGDSGGPLVSSDARDIWYLAG
IVSWGDECAKPNKPGVYTRVTALRDWITSKTI

Transmembrane domain:

amino acids 21-40 (type II)

FIGURE 107

AGAGAAAGAAGCGTCTCCAGCTGAAGCCAATGCAGCCCTCCGGCTCTCCGCGAAGAAGTTCCTG
 CCCCAGATGAGCCCCCGCGTCCCGACTATCCCCAGGCGGGCGTGGGGCACCGGGCCCAGC
 GCCGACGATCGCTGCCGTTTTGCCCTTGGGAGTAGGATGTGGTGAAAGGATGGGGCTTCTCCCTT
 ACGGGGCTCACAAATGGCCAGAGAAGATTCCGTGAAGTGTCTGCGCTGCCTGCTCTACGCCCTCAA
 TCTGCTCTTTTGGTTAATGTCCATCAGTGTGTTGGCAGTTTCTGCTTGGATGAGGGACTACCTAA
 ATAATGTTCTCACTTTAACTGCAGAAACGAGGGTAGAGGAAGCAGTCATTTGACTTACTTTCCCT
 GTGGTTCATCCGTCATGATTGCTGTTTGTCTTTTCTTATCATTGTGGGGATGTTAGGATATTG
 TGGAACGGTGAAAGAAATCTGTTGCTTCTTGCATGGTACTTTGGAAGTTTGTGTCATTTTCT
 GTGTAGAAGTGGCTTGTGGCGTTTGGACATATGAACAGGAAGTATGGTTCCAGTACAATGGTCA
 GATATGGTCACTTTGAAGCCAGGATGACAAATTATGGATTACCTAGATATCGGTGGCTTACTCA
 TGCTTGGAAATTTTTTTCAGAGAGAGTTTAAAGTGTCTGTGGAGTAGTATATTTCACTGAGTGGTTGG
 AAATGACAGAGATGGACTGGCCCCAGATTCTGCTGTGTTAGAGAATCCAGGATGTTCCAAA
 CAGGCCCCACCAGGAAGATCTCAGTGACCTTTATCAAGAGGGTGTGGGAAGAAAATGTATTCCTT
 TTTGAGAGGAACCAACAACACTGCAGGTGCTGAGGTTTCTGGGAATCTCCATTGGGGTGACACAAA
 TCCTGGCCATGATTCTCACCATTACTCTGCTCTGGGCTCTGTATTATGATAGAAGGGAGCCTGGG
 ACAGACCAATGATGTCCTTGAAGAATGACAACTCTCAGCACCTGTCTGCTCCCTCAGTAGAACT
 GTTGAACCAAGCCTGTCAAGAATCTTTGAACACACATCCATGGCAAACAGCTTTAATACACACT
 TTGAGATGGAGGAGTTATAAAAGAAATGTCACAGAAGAAAACCAAACTTGTTTTATTGGACT
 TGTGAATTTTTGAGTACATACTATGTGTTTCAGAAATATGTAGAAATAAAAATGTTGCCATAAAA
 TAACACCTAAGCATATACTATTCTATGCTTTAAAATGAGGATGGAAAAGTTTCATGTCATAAGTC
 ACCACCTGGACAATAATTGATGCCCTTAAAATGCTGAAGACAGATGTCATACCCACTGTGTAGCC
 TGTGTATGACTTTTACTGAACACAGTTATGTTTTGAGGCAGCATGGTTTGATTAGCATTTCGCCA
 TCCATGCAAACGAGTCACATATGGTGGGACTGGAGCCATAGTAAAGGTTGATTACTTCTACCAA
 CTAGTATATAAAGTACTAATTAAATGCTAACATAGGAAGTTAGAAAATACTAATAACTTTTATTA
 CTCAGCGATCTATTCTTCTGATGCTAAATAAATATATATCAGAAAACCTTCAATATTGGTGACT
 ACCTAAATGTGATTTTTGCTGGTTACTAAAATATCTTACCACCTAAAAGAGCAAGCTAACACAT
 TGTCTTAAGCTGATCAGGGATTTTTGTATATAAGTCTGTGTAAATCTGTATAATTCAGTCGAT
 TTCAGTTCTGATAATGTTAAGAATAACCATTATGAAAAGGAAAATTTGTCCTGTATAGCATCATT
 ATTTTATAGCCTTTCTGTTAATAAAGCTTTACTATTCTGTCTGGGCTTATATTACACATATAAC
 TGTTATTTAAATACTTAACCCTAATTTTGAATAATTACCAGTGTGATACATAGGAATCATTTATTC
 AGAATGTAGTCTGGTCTTTAGGAAGTATTAATAAGAAAATTTGCACATAACTTAGTTGATTGAGA
 AAGGACTTGTATGCTGTTTTTCTCCAAATGAAGACTCTTTTGTACACTAAACACTTTTTAAAAA
 GCTTATCTTTGCCCTTCTCAAACAAGAAGCAATAGTCTCCAAGTCAATATAAATTCTACAGAAAA
 TAGTGTCTTTTTTCTCCAGAAAAATGCTTGTGAGAATCATTTAAACATGTGACAATTTAGAGATT
 CTTTGTTTTATTTCACTGATTAATATACTGTGGCAAATTACACAGATTATTAAATTTTTTTACAA
 GAGTATAGTATATTTATTTGAAATGGGAAAAGTGCATTTTACTGTATTTTGTGATTTTGTGTTAT
 TTCTCAGAATATGGAAAGAAAATTAAAATGTGTCAATAAATATTTTCTAGAGAGTAA

FIGURE 108

MAREDSVKCLRCLLYALNLLFWLMSISVLAVSAWMRDYLNNVLTTLTAETRVEEAVILTYFPVVHP
VMIAVCCFLIIVGMLGYCGTVKRNLLLLLAWYFGSLLVIFCVELACGVWTYEQELMVPVQWSDMT
LKARMTNYGLPRYRWLTHAWNFFQREFKCCGVVYFTDWLEMTMDWPPDSCCVREFFPGCSKQAHQ
EDLSDLYQEGCGKMYSFRLGTKQLQVLRFLGISIGVTQILAMILTITLLWALYYDRREPQTDQM
MSLKNDNSQHLSCPSVELLKPSLSRIFEHTSMANSFNTHFEMEEL

Signal peptide:

amino acids 1-33

Transmembrane domains:

amino acids 12-35, 57-86, 94-114, 226-248

FIGURE 109

CCAAGGCCAGAGCTGTGGACACCTTATCCCACTCATCCTCATCCTCTTCTCTGATAAAGCCCCCTACCAGTGCT
GATAAAGTCTTTCTCGTGAGAGCCTAGAGGCCCTTAAAAAAGTGTGAAAGAGAAGGGGACAAAGGAACA
CCAGTATTAAGAGGATTTTCCAGTGTCTTGGCAGTTGGTCCAGAAGGATGCTCCATTCTGCTTCTCACCTG
CCTCTTCATCAGGACCTCCGTGTACCCGTGGCCCTAGATCCTTGTCTGCTTACATCAGCCTGAATGAGC
CCTGGAGGAACACTGACCACCAGTTGGATGAGTCTCAAGGTCTCCTCTATGTGACAACCATGTGAATGGGAG
TGGTACCCTTACGGGCATGGCGGGAGATGCCATGCCTACCTTCTGCATACCAGAAAACCACTGTGGAACCCA
CGCACCTGTCTGGCTCAATGGCAGCCACCCCTAGAAGCGACGGCATTGTGCAACGCCAGGCTTGTGCCAGCT
TCAATGGGAAGTGTGTCTCTGGAACACCACGGTGAAGTCAAGGCTTGCCCTGGAGGCTACTATGTGTATCGT
CTGACCAAGCCAGCGTCTGCTTCCACGTCTACTGTGGTCATTTTTATGACATCTGCGACGAGGACTGCCATGG
CAGCTGCTCAGATACCAGCGAGTGCACATGCGCTCCAGGAAGTGTGCTAGGCCCTGACAGGCAGACATGCTTTG
ATGAAAATGAATGTGAGCAAAACAACGGTGGCTGCAGTGAGATCTGTGTGAACCTCAAAACTCCTACCCTGT
GAGTGTGGGGTGGCCGTGTGCTAAGAAGTGATGGCAAGACTTGTGAAGACGTTGAAGGATGCCACAATAACAA
TGGTGGCTGCAGCCACTCTTGCTTGGATCTGAGAAAGGCTACCAGTGTGAATGTCCCCGGGGCTGGTGTGT
CTGAGGATAACCACTTGCCAAGTCCCTGTGTTGTGCAAATCAAATGCCATTGAAGTGAACATCCCCAGGGAG
CTGTTTGGTGGCTGGAGCTCTTCCCTGACCAACACCTCCTGCCGAGGAGTGTCCAACGGCAGCCCATGTCAACAT
CCTCTTCTCTCAAGACATGTGGTACAGTGGTGCATGTGGTGAATGACAAGATTGTGGCCAGCAACCTCGTGA
CAGTCTTACCAAGCAGACCCCGGGGAGCAGCGGGGACTTCATCATCCGAACCAGCAAGCTGCTGATCCCGGTG
ACCTGCGAGTTTCCACGCTGTACACCATTCTGAAGGATACGTTCCCAACCTTCGAACTCCCCACTGGAAAT
CATGAGCCGAAATCATGGGATCTTCCCATCTACTGAGATCTTCAAGGACAATGAGTTTGAAGAGCCTTACC
GGGAAGCTCTGCCACCCCTCAAGCTTCGTGACTCCCTCTACTTTGGCATTGAGCCCGTGGTGCACGTGAGCGGC
TTGGAAAGCTTGGTGGAGAGCTGCTTTGCCACCCACCTCCAAGATCGACGAGGTCTGAAATACTACCTCAT
CCGGGATGGCTGTGTTTCAGATGACTCGGTAAAGCAGTACACATCCCGGGATCACCTAGCAAAGCACTTCCAGG
TCCCTGTCTTCAAGTTTGTGGGCAAAGACCACAAGGAAGTGTCTGCACTGCCGGGTTCTTGTCTGTGGAGTG
TTGGACGAGCGTTCCCGCTGTGCCAGGGTTGCCACCGGCAATGCGTCGTGGGGCAGGAGGAGAGGACTCAGC
CGGTCTACAGGGCCAGACGCTAACAGGCGGCCGATCCGCATCGACTGGGAGGACTAGTTCTGTAGCCATACCTC
GAGTCCCCTGCATTGGACGGCTCTGCTCTTTGGAGCTTCTCCCCCACC GCCCTCTAAGAACATCTGCCAACAGC
TGGGTTTCACTTACACTGTGAGTTCACTCCCAGCACCACCTCACTCTGATTCTGGTCCATTCACTGGGCA
CAGGTACAGCACTGCTGAACAATGTGGCCTGGGTGGGTTTCATCTTCTAGGGTTGAAACTAACTGTCCA
CCCAGAAAGACACTCACCCATTTCCCTCATTTCTTCTACACTTAAATACCTCGTGTATGGTGCAATCAGAC
CACAAAATCAGAAGCTGGGTATAATATTTCAAGTTACAAACCCTAGAAAAATTAAACAGTTACTGAAATTATGA
CTTAAATACCAATGACTCCTTAAATATGTAAATTATAGTTATACCTTGAAATTTCAATTCAAATGCAGACTAA
TTATAGGGAATTTGGAAGTGATCAATAAAACAGTATATAATTTT

FIGURE 110

MPPFLLLTCLFITGTSVSPVALDPCSAYISLNEPWRNTDHLDESQGPPLCDNHVNGEWYHFTGMAGDAMP
TFCIPENHCGTHAPVWLNGSHPLEGDGIVQRQACASFNGNCLWNTTVEVKACPGGYVYRLTKPSVCFHV
YCGHFYDIDCEDCHGSCSDTSECTCAPGTVLGPDRQTCFDENECEQNNGGCSEICVNLKNSYRCECGVGRV
LRSDGKTCEDVEGCHNNNGGCSSHCLGSEKGYQCECPRGLVLSEDNHTCQVPVLCKSNAIEVNIPRELVGG
LELFLTNTSCRGVSNNGTHVNILFSLKTCGTVVDVNDKIVASNLVTGLPKQTPGSSGDFIIRTSKLLIPVT
CEFPRLYTISEGYVPNLRNSPLEIMSRNHGIFPFTLEIFKDNEFEEPYREALPTLKLRLDSLYFGIEPVVHV
SGLESLVESCFATPTSKIDEVLKYILIRDGCVSDDSVKQYTSRDHLAKHFQVPVFKFVKGDKHEVFLHCRV
LVCGVLDERSRCAQGCHRRMRRGAGGEDSAGLQGQTLTGGPPIRIDWED

Important features of the protein:

Signal peptide:

amino acids 1-16

N-glycosylation sites.

amino acids 89-93, 116-120, 259-263, 291-295, 299-303

Tyrosine kinase phosphorylation sites.

amino acids 411-418, 443-451

N-myristoylation sites.

amino acids 226-232, 233-239, 240-246, 252-258, 296-302, 300-306,
522-528, 531-537

Aspartic acid and asparagine hydroxylation site.

amino acids 197-209

ZP domain proteins.

amino acids 431-457

Calcium-binding EGF-like proteins.

amino acids 191-212, 232-253

FIGURE 111

GAGAGAGGCAGCAGCTTGCTCAGCGGACAAGGATGCTGGGCGTGAGGGACCAAGGCCTGCCCTGCACTCGG
GCCTCCTCCAGCCAGTGTCTGACCAGGGACTTCTGACCTGCTGGCCAGCCAGGACCTGTGTGGGGAGGCCCT
CCTGCTGCCTTGGGGTGACAATCTCAGCTCCAGGCTACAGGGAGACCGGGAGGATCACAGAGCCAGCATGT
TACAGGATCCTGACAGTGATCAACCTCTGAACAGCCTCGATGTCAAACCCCTGCGCAAACCCCGTATCCCC
ATGGAGACCTTCAGAAAGGTGGGGATCCCCATCATCATAGCACTACTGAGCCTGGCGAGTATCATCATTGT
GGTTGTCTCATCAAGGTGATTCTGGATAAATACTACTTCTCTGCGGGCAGCCTCTCCACTTCATCCCGA
GGAAGCAGCTGTGTGACGGAGAGCTGGACTGTCCCTTGGGGGAGGACGAGGAGCACTGTGTCAAGAGCTTC
CCCGAAGGGCCTGCAGTGGCAGTCCGCCCTCTCCAAGGACCGATCCACACTGCAGGTGCTGGACTCGGCCAC
AGGGAAGTGGTTCTCTGCCTGTTTCGACAACCTTCACAGAAGCTCTCGCTGAGACAGCCTGTAGGCAGATGG
GCTACAGCAGAGCTGTGGAGATTGGCCCAGACCAGGATCTGGATGTTGTTGAAATCACAGAAAACAGCCAG
GAGCTTCGCATGCGGAACCTCAAGTGGGCCCTGTCTCTCAGGCTCCCTGGTCTCCCTGCACTGTCTTGCCTG
TGGGAAGAGCCTGAAGACCCCCCGTGTGGTGGGTGGGGAGGAGGCCTCTGTGGATTCTTGGCCTTGGCAGG
TCAGCATCCAGTACGACAAACAGCACGTCTGTGGAGGGAGCATCCTGGACCCCCACTGGGTCTCAGGGCA
GCCCCACTGCTTCAGGAAACATACCGATGTGTTCAACTGGAAGGTGCGGGCAGGCTCAGACAAACTGGGCAG
CTTCCCCATCCCTGGCTGTGGCCAAGATCATCATATTGAATTCAACCCCATGTACCCCAAAGACAATGACA
TCGCCCTCATGAAGCTGCAGTTCCCACTCACTTTCTCAGGCACAGTCAGGCCCATCTGTCTGCCCTTCTTT
GATGAGGAGCTCACTCCAGCCACCCCACTCTGGATCATTTGATGGGGCTTTACGAAGCAGAATGGAGGGAA
GATGTCTGACATACTGCTGCAGGCGTCAGTCCAGGTCATTGACAGCACACGGTGCAATGCAGACGATGCGT
ACCAGGGGGGAAGTCACCGAGAAGATGATGTGTGCAGGCATCCCGGAAGGGGTGTGGACACCTGCCAGGGT
GACAGTGGTGGGCCCTGATGTACCAATCTGACCAGTGGCATGTGGTGGGCATCGTTAGCTGGGGCTATGG
CTGCGGGGGCCCGAGCACCCCAAGGATATACACCAAGGTCTCAGCCTATCTCAACTGGATCTACAATGTCT
GGAAGGCTGAGCTGTAATGCTGCTGCCCCCTTGCACTGCTGGGAGCCGCTTCCTTCTGCCCCTGCCACCT
GGGGATCCCCCAAAGTCAGACACAGAGCAAGAGTCCCTTGGGTACACCCCTCTGCCACAGCCTCAGCAT
TTCTTGGAGCAGCAAAGGGCCTCAATTCTGTGAAGAGACCTCGCAGCCCAGAGCGCCCAGAGGAAGTCA
GCAGCCCTAGCTCGGCCACACTTGGTGCTCCCAGCATCCCAGGGAGAGACACAGCCCACTGAACAAGGTCT
CAGGGGTATTGCTAAGCCAAGAAGGAACCTTCCCACTACTGAATGGAAGCAGGCTGTCTTGTAAGGCC
CAGATCACTGTGGGCTGGAGAGGAGAAGGAAAGGGTCTGCGCCAGCCCTGTCCGTCTTACCCATCCCCAA
GCCTACTAGAGCAAGAAACAGTTGTAATATAAAATGCACTGCCCTACTGTTGGTATGACTACCGTTACCT
ACTGTTGTCATTGTTATTACAGCTATGGCCACTATTATTAAAGAGCTGTGTAAACATCTCTGGCAAAAAAA
AAAA

FIGURE 112

MLQDPDSDQPLNSLDVKPLRKPRIPMETFRKVGIPIIIIALLSLASIIIVVVLIKVILDKYYFLCG
QPLHFIPRKQLCDGELDCPLGEDEEHCVKSFPEGPAVAVRLSKDRSTLQVLDSATGNWFSACFDN
FTEALAETACRQMGYSRAVEIGPDQDLVDVEITENSQELMRNSSGPCLSGSLVSLHCLACGKSL
KTPRVVGEEASVDSWPWQVSIQYDKQHVCSSILDPHWVLTAAHCFRKHTDVFNWKVRAGSDKL
GSFPSLAVAKIIIEFNPMYPKDNDIALMKLQFPLTFSGTVRPICLPFFDEELTPATPLWIIIGWG
FTKQNGGKMSDILLQASVQVIDSTRCNADDAYQGEVTEKMMCAGIPEGGVDTCCQGDSSGGLMYQS
DQWHVVGIVSWGYGCGGPSTPGVYTKVSAYLNWIYNVWKAEL

Transmembrane domain:

amino acids 32-53 (typeII)

FIGURE 113

GGCTGGACTGGAACCTCTGGTCCCAAGTGATCCACCCGCTCAGCCTCCCAAGGTGCTGTGATTA
TAGGTGTAAGCCACCGTGTCTGGCCTCTGAACAACCTTTTTCAGCAACTAAAAAGCCACAGGAGT
TGAAGTGTAGGATTCTGACTATGCTGTGGTGGCTAGTGCTCCTACTCCTACCTACATTAAATC
TGTTTTTTGTTCTCTTGTAAGTACCTTTACCTTCCTAACACAGAGGATCTGTCACTGTGGCTCT
GGCCCAAACCTGACCTTCACCTCTGGAACGAGAACAGAGGTTTCTACCCACACCGTCCCTCGAAG
CCGGGGACAGCCTCACCTTGCTGGCCTCTCGCTGGAGCAGTGCCCTCACCAACTGTCTCACGTCT
GGAGGCACTGACTCGGGCAGTGCAGGTAGCTGAGCCTCTTGGTAGCTGCGGCTTTCAGGTGGGC
CTTGCCCTGGCCGTAGAAGGGATTGACAAGCCCGAAGATTTTCATAGGCGATGGCTCCCACTGCCC
AGGCATCAGCCTTGCTGTAGTCAATCACTGCCCTGGGGCCAGGACGGGCCGTGGACACCTGCTCA
GAAGCAGTGGGTGAGACATCACGCTGCCCGCCCATCTAACCTTTTCATGTCCTGCACATCACCTG
ATCCATGGGCTAATCTGAAGTCTGTCCCAAGGAACCCAGAGCTTGAGTGAGCTGTGGCTCAGACC
CAGAAGGGGTCTGCTTAGACCACCTGGTTTTATGTGACAGGACTTGCAATTCTCCTGGAACATGAGG
GAACGCCGGAGGAAAGCAAAGTGGCAGGGAAGGAACCTTGTCCAAATTATGGGTGAGAAAAGATG
GAGGTGTTGGGTTATCACAAGGCATCGAGTCTCCTGCATTCACTGGACATGTGGGGGAAGGGCTG
CCGATGGCGCATGACACACTCGGGACTCACCTCTGGGGCCATCAGACAGCCGTTTCCGCCCCGAT
CCACGTACCAGCTGCTGAAGGGCAACTGCAGGCCGATGCTCTCATCAGCCAGGCAGCAGCCAAAA
TCTGCGATCACCAGCCAGGGGCAGCCGTCTGGGAAGGAGCAAGCAAAGTGACCATTTCTCCTCCC
CTCCTTCCCTCTGAGAGGCCCTCCTATGTCCCTACTAAAGCCACCAGCAAGACATAGCTGACAGG
GGCTAATGGCTCAGTGTTGGCCCAGGAGGTGAGCAAGGCCTGAGAGCTGATCAGAAGGGCCTGCT
GTGCGAACACGGAAATGCCTCCAGTAAGCACAGGCTGCAAAATCCCCAGGCAAAGGACTGTGTGG
CTCAATTTAAATCATGTTCTAGTAATTGGAGCTGTCCCCAAGACCAAAGGAGCTAGAGCTTGTT
CAAATGATCTCCAAGGGCCCTTATACCCAGGAGACTTTGATTTGAATTTGAAACCCCAAATCCA
AACCTAAGAACCAGGTGCATTAAAGAATCAGTTATTGCCGGGTGTGGTGGCCTGTAATGCCAACAT
TTTGGGAGGCCGAGGCGGGTAGATCACCTGAGGTGAGGAGTTCAAGACCAGCCTGGCCAAATGG
TGAAACCCCTGTCTCTACTAAAAATACAAAAAACTAGCCAGGCATGGTGGTGTGTGCCTGTATC
CCAGCTACTCGGGAGGCTGAGACAGGAGAATTACTTGAACCTGGGAGGTGAAGGAGGCTGAGACA
GGAGAATCACTTCAGCCTGAGCAACACAGCGAGACTCTGTCTCAGAAAAATAAAAAAGAATTA
TGGTTATTTGTAA

FIGURE 114

MLWWLVLLLLPTLKSVCSLVTSLYLPNTEDLSLWLWPKPDLHSGTRTEVSTHTVPSKPGTASPC
WPIAGAVPSPTVSRLEALTRAVQVAEPLGSCGFGGPCPGRRRD

Signal peptide:
amino acids 1-15

FIGURE 115

CAGCAGTGGTCTCTCAGTCCTCTCAAAGCAAGGAAAGAGTACTGTGTGCTGAGAGACCATGGCAA
AGAATCCTCCAGAGAATTGTGAAGACTGTCACATTCTAAATGCAGAAGCTTTTAAATCCAAGAAA
ATATGTAAATCACTTAAGATTTGTGGACTGGTGTGGTATCCTGGCCCTAACTCTAATTGTCCT
GTTTTGGGGGAGCAAGCACTTCTGGCCGGAGGTACCCAAAAAGCCTATGACATGGAGCACACTT
TCTACAGCAATGGAGAGAAGAAGAAGATTTACATGGAAATTGATCCTGTGACCAGAACTGAAATA
TTCAGAAGCGGAAATGGCACTGATGAAACATTGGAAGTGCACGACTTTAAAAACGGATACACTGG
CATCTACTTCGTGGGTCTTCAAAAATGTTTTATCAAACTCAGATTAAAGTGATTCCTGAATTTT
CTGAACCAGAAGAGGAAATAGATGAGAATGAAGAAATTACCACAACCTTCTTTGAACAGTCAGTG
ATTTGGGTCCCAGCAGAAAAGCCTATTGAAAACCGAGATTTCTTAAAAATTCCAAAATTCTGGA
GATTTGTGATAACGTGACCATGTATTGGATCAATCCCACTCTAATATCAGTTTCTGAGTTACAAG
ACTTTGAGGAGGAGGAGAAGATCTTCACTTTCCTGCCAACGAAAAAAGGGATTGAACAAAAT
GAACAGTGGGTGGTCCCTCAAGTGAAAGTAGAGAAGACCCGTCACGCCAGACAAGCAAGTGAGGA
AGAACTTCCAATAAATGACTATACTGAAAATGGAATAGAATTTGATCCCATGCTGGATGAGAGAG
GTTATTGTTGTATTTACTGCCGTCGAGGCAACCGCTATTGCCGCCGCTCTGTGAACCTTTACTA
GGCTACTACCCATATCCATACTGCTACCAAGGAGGACGAGTCATCTGTGCTGTCATCATGCCTTG
TAACTGGTGGGTGGCCCGCATGCTGGGGAGGGTCTTAATAGGAGGTTTGAGCTCAAATGCTTAAAC
TGCTGGCAACATATAATAAATGCATGCTATTCAATGAATTTCTGCCTATGAGGCATCTGGCCCT
GGTAGCCAGCTCTCCAGAATTACTTGTAGGTAATTCCTCTCTTCATGTTCTAATAAACTTCTACA
TTATCACCAAAAAAAAAAAAAAAAAAAAA

FIGURE 116

MAKNPPENCEDCHILNAEAFKSKKICKSLKICGLVFGILALTLIVLFWGSKHFWPEVPPKAYDME
HTFYSNGEKKKIYMEIDPVTRTEIFRSGNGTDETLEVHDFKNGYTGIIYFVGLQKCFIKTQIKVIP
EFSEPEEEIDENEEITTTFFEQSVIWWPAEKPIENRDFLKNKILEICDNVTMYWINPTLISVSE
LQDFEEEGEDLHFPANEKKGIEQNEQWVVPQVKVEKTRHARQASEEELPINDYTENGIEFDPMLD
ERGYCCIIYCRRGNRYCRRVCEPLLGYYPYPYCYQGGRVICRVIMPCNWWVARMLGRV

Important features of the protein:

Signal peptide:

amino acids 1-40

Transmembrane domain:

amino acids 25-47 (type II)

N-glycosylation sites.

amino acids 94-97, 180-183

Glycosaminoglycan attachment sites.

amino acids 92-95, 70-73, 85-88, 133-136, 148-151, 192-195, 239-
242

N-myristoylation sites.

amino acids 33-38, 95-100, 116-121, 215-220, 272-277

Microbodies C-terminal targeting signal.

amino acids 315-317

Cytochrome c family heme-binding site signature.

amino acids 9-14

FIGURE 117

GAGCTCCCCCTCAGGAGCGCGTTAGCTTCACACCTTCGGCAGCAGGAGGGCGGCAGCTTCTCGCAGGCGGCA
GGGCGGGCGGCCAGGATCATGTCCACCACCACATGCCAAGTGGTGGCGTTCCTCCTGTCCATCCTGGGGCT
GGCCGGCTGCATCGCGGCCACCGGGATGGACATGTGGAGCACCAGGACCTGTACGACAACCCCGTCACCT
CCGTGTTCCAGTACGAAGGGCTCTGGAGGAGCTGCGTGAGGCAGAGTTCAGGCTTCACCGAATGCAGGCCC
TATTTACCATCCTGGGACTTCCAGCCATGCTGCAGGCAGTGCAGGCCCTGATGATCGTAGGCATCGTCTT
GGGTGCCATTGGCCTCCTGGTATCCATCTTTGCCCTGAAATGCATCCGCATTGGCAGCATGGAGGACTCTG
CCAAAGCCAACATGACACTGACCTCCGGGATCATGTTTCATTGTCTCAGGTCTTTGTGCAATTGCTGGAGTG
TCTGTGTTTGCCAACATGCTGGTACTAAGTCTGGATGTCCACAGCTAACATGTACACCGGCATGGGTGG
GATGGTGCAGACTGTTTACAGACAGGTACACATTTGGTGCGGCTCTGTTCTGGGGCTGGGTGCTGGAGGCC
TCACACTAATTGGGGGTGTGATGATGTGCATCGCCTGCCGGGGCCTGGCACCAGAAGAAACCAACTACAAA
GCCGTTTCTTATCATGCCCTCAGGCCACAGTGTGCTTACAAGCCTGGAGGCTTCAAGGCCAGCACTGGCTT
TGGGTCCAAACCAAAAAACAAGAATATACGATGGAGGTGCCCGCACAGAGGACGAGGTACAATCTTATC
CTTCCAAGCACGACTATGTTAATGCTCTAAGACCTCTCAGCACGGGCGGAAGAACTCCCGGAGAGCTCA
CCCCAAAAACAAGGAGATCCCATCTAGATTTCTTCTGCTTTTGACTCACAGCTGGAAGTTAGAAAAGCCT
CGATTTTCATCTTTGGAGAGGCCAAATGGTCTTAGCCTCAGTCTCTGTCTCTAAATATTCCACCATAAAACA
GCTGAGTTATTTATGAATTAGAGGCTATAGCTCACATTTTCAATCTCTATTTCTTTTTTTAAATATAACT
TTCTACTCTGATGAGAGAATGTGGTTTTAATCTCTCTCTCACATTTTGATGATTTAGACAGACTCCCCCTC
TTCCTCCTAGTCAATAAACCATTGATGATCTATTTCCAGCTTATCCCCAAGAAAACCTTTTGAAAGGAAA
GAGTAGACCCAAAGATGTTATTTCTGCTGTTTGAATTTTGTCTCCCCACCCCAACTTGGCTAGTAATAA
ACACTTACTGAAGAAGAAGCAATAAGAGAAAGATATTTGTAATCTCTCCAGCCCATGATCTCGGTTTTCTT
ACACTGTGATCTTAAAGTTACCAAAACCAAGTCATTTTCAGTTTGAGGCAACCAACCTTTCTACTGCTG
TTGACATCTTCTTATTACAGCAACACCATTCTAGGAGTTTCCTGAGCTCTCCACTGGAGTCTCTTTCTGT
CGCGGGTCAGAAATTGTCCCTAGATGAATGAGAAAATTATTTTTTTAATTTAAGTCCTAAATATAGTTAA
AATAAATAATGTTTTAGTAAATGATACACTATCTCTGTGAAATAGCCTCACCCTACATGTGGATAGAAG
GAAATGAAAAAATAATTGCTTTGACATTGTCTATATGGTACTTTGTAAAGTCATGCTTAAGTACAAATTC
ATGAAAAGCTCACACCTGTAATCCTAGCACTTTGGGAGGCTGAGGAGGAAGGATCACTTGAGCCCAGAAGT
TCGAGACTAGCCTGGGCAACATGGAGAAGCCCTGTCTCTACAAAATACAGAGAGAAAAATCAGCCAGTCA
TGGTGGCATACACCTGTAGTCCCAGCATTCGGGAGGCTGAGGTGGGAGGATCACTTGAGCCCAGGGAGGT
TGGGCTGCAGTGAGCCATGATCACACCACTGCCTCCAGCCAGGTGACATAGCGAGATCCTGTCTAAAAA
AATAAAAAATAAATAATGGAACACAGCAAGTCCTAGGAAGTAGGTTAAACTAATTCCTTAA

FIGURE 118

MSTTTCQVAFLLSILGLAGCIAATGMDMWSTQDLYDNPVTSVVFQYEGLWRSCVRQSSGFTECRP
YFTILGLPAMLQAVRALMIVGIVLGAIGLLVSIFALKCIRIGSMEDSAKANMTLTSGIMFIVSGL
CAIAGVSVFANMLVTNFWMSTANMYTGMGGMVQTVQTRYTFGAALFVGWVAGGLTLIGGVMMCIA
CRGLAPEETNYKAVSYHASGHSVAYKPGGFKASTGFGSNTKNKKIYDGGARTEDEVQSYPSKHDY
V

Signal peptide:

amino acids 1-23

Transmembrane domains:

amino acids 81-100, 121-141, 173-194

FIGURE 119

GGAAAACTGTTCTCTTCTGTGGCACAGAGAACCCTGCTTCAAAGCAGAAGTAGCAGTTCGGGAGTCC
AGCTGGCTAAAACTCATCCCAGAGGATAATGGCAACCCATGCCTTAGAAATCGCTGGGCTGTTTCTTG
GTGGTGTGGGAATGGTGGGCACAGTGGCTGTCACTGTCATGCCTCAGTGGAGAGTGTGGCCTTCATT
GAAAACAACATCGTGGTTTTTGAAAACCTCTGGGAAGGACTGTGGATGAATTGCGTGAGGCAGGCTAA
CATCAGGATGCAGTGCAAAATCTATGATCCCTGCTGGCTCTTCTCCGGACCTACAGGCAGCCAGAG
GACTGATGTGTGCTGCTTCCGTGATGTCTTCTTGGCTTTCATGATGGCCATCCTTGGCATGAAATGC
ACCAGGTGCACGGGGGACAATGAGAAGGTGAAGGCTCACATTCTGCTGACGGCTGGAATCATCTTCAT
CATCACGGGCATGGTGGTGTCTATCCCTGTGAGCTGGGTTGCCAATGCCATCATCAGAGATTTCTATA
ACTCAATAGTGAATGTTGCCAAAAACGTGAGCTTGGAGAAGCTCTCTACTTAGGATGGACCACGGCA
CTGGTGTGCTGATTGTTGGAGGAGCTCTGTTCTGCTGCGTTTTTTTGTGCAACGAAAAGAGCAGTAGCTA
CAGATACTCGATACCTTCCCATCGCACAAACCAAAAAAGTTATCACACCGAAAGAAGTCACCGAGCG
TCTACTCCAGAAGTCAGTATGTGTAGTTGTGTATGTTTTTTAACTTTACTATAAAGCCATGCAAATG
ACAAAAATCTATATTACTTTCTCAAAATGGACCCCAAGAACTTTGATTTACTGTTCTTAACTGCCT
AATCTTAATTACAGGAAGTGTGCATCAGCTATTTATGATTCTATAAGCTATTTACAGAGAATGAGATA
TTAAACCCAAATGCTTTGATTGTTCTAGAAAGTATAGTAATTTGTTTTCTAAGGTGGTTCAAGCATCTA
CTCTTTTTATCATTTACTTCAAATGACATTGCTAAAGACTGCATTATTTTACTACTGTAATTTCTCC
ACGACATAGCATTATGTACATAGATGAGTGTAAACATTTATATCTCACATAGAGACATGCTTATATGGT
TTTATTTAAATGAAATGCCAGTCCATTACACTGAATAAATAGAACTCAACTATTGCTTTTCAGGGAA
ATCATGGATAGGGTTGAAGAAGGTACTATTAATTGTTTAAAAACAGCTTAGGGATTAATGTCTCCCA
TTTATAATGAAGATTAAATGAAGGCTTTAATCAGCATTGTAAAGGAAATTGAATGGCTTTCTGATAT
GCTGTTTTTTAGCCTAGGAGTTAGAAATCCTAACTTCTTTATCCTCTTCTCCAGAGGCTTTTTTTTT
CTTGTGTATTAAATTAACATTTTTTAAACGCAGATATTTTGTCAAGGGGCTTTGCATTCAAACGCTT
TTCCAGGGCTATACTCAGAAGAAAGATAAAAGTGTGATCTAAGAAAAAGTGATGGTTTTAGGAAAGTG
AAAATATTTTTGTTTTTGTATTTGAAGAAGATGATGCATTTTGACAAGAAATCATATATGTATGGAT
ATATTTTAATAAGTATTTGAGTACAGACTTTGAGGTTTCATCAATATAAATAAAGAGCAGAAAAATA
TGTCTTGGTTTTTCATTTGCTTACCAAAAAACAACAACAAAAAAGTTGTCCTTTGAGAACTTCACCT
GCTCCTATGTGGGTACCTGAGTCAAAATTGTCATTTTTGTTCTGTGAAAAATAAATTCCTTCTTGTA
CCATTTCTGTTTAGTTTTACTAAAACTGTAAATACTGTATTTTCTGTTTATTCCAAATTTGATGAA
ACTGACAATCCAATTTGAAAGTTTGTGTGACGTCTGTCTAGCTTAAATGAATGTGTTCTATTGCTT
TATACATTTATATTAATAAATTGTACATTTTTTCTAATT

FIGURE 120

MATHALEIAGLFLGGVGMVGTVAVTVMPQWRVSAFIENNIVVFENFWGLWMNCVRQANIRMQCK
IYDSLALLSPDLQAARGLMCAASVMSFLAFMAILGMKCTRCTGDNEKVKAHILLTAGIIFIITG
MVLIPVSWVANAIIRDFYNSIVNVAQKRELGEALYLGWTTALVLIVGGALFCCVFCCNEKSSSY
RYSIPSHRTTQKSYHTGKKSPSVYSRSQYV

Signal peptide:

amino acids 1-17

Transmembrane domains:

amino acids 82-101, 118-145, 164-188

FIGURE 121

GGAGAGAGGCGCGCGGGTGAAAGGCGCATTGATGCAGCCTGCGGGCGGCCTCGGAGCGCGGCGGAG
CCAGACGCTGACCACGTTCCCTCTCCTCGGTCTCCTCCGCCTCCAGCTCCGCGCTGCCCCGCGAGCC
GGGAGCCATGCGACCCCAGGGCCCCGCGCCTCCCCGCGAGCGGCTCCGCGGCCTCCTGCTGCTCC
TGCTGCTGCAGCTGCCCCGCGCGTTCGAGCGCCTCTGAGATCCCCAAGGGGAAGCAAAAGGCGCAG
CTCCGGCAGAGGGAGGTGGTGGACCTGTATAATGGAATGTGCTTACAAGGGCCAGCAGGAGTGCC
TGGTCGAGACGGGAGCCCTGGGGCCAATGTTATTCCGGGTACACCTGGGATCCCAGGTCGGGATG
GATTCAAAGGAGAAAAGGGGAATGTCTGAGGGAAAGCTTTGAGGAGTCCTGGACACCCAACCTAC
AAGCAGTGTTTCATGGAGTTCATTGAATTATGGCATAGATCTTGGGAAAATTGCGGAGTGATACATT
TACAAAGATGCGTTCAAATAGTGCTCTAAGAGTTTTGTTTCAGTGGCTCACTTCGGCTAAAATGCA
GAAATGCATGCTGTGAGCGTTGGTATTTACATTCAATGGAGCTGAATGTTTCAGGACCTCTTCCC
ATTGAAGCTATAATTTATTTGGACCAAGGAAGCCCTGAAATGAATTCAACAATTAATATTCATCG
CACTTCTTCTGTGGAAGGACTTTGTGAAGGAATTGGTGTGCTGGATTAGTGGATGTTGTCTATCTGGG
TTGGCACTTGTTTCAGATTACCCAAAAGGAGATGCTTCTACTGGATGGAATTCAGTTTCTCGCATC
ATTATTGAAGAACTACCAAAATAAATGCTTTAATTTTCATTGCTACCTCTTTTTTTATTATGCC
TTGGAATGGTTCACTTAAATGACATTTTAAATAAGTTTATGTATACATCTGAATGAAAAGCAAAG
CTAAATATGTTTACAGACCAAAGTGTGATTTACACTGTTTTTAAATCTAGCATTATTCATTTTG
CTTCAATCAAAGTGGTTTCAATATTTTTTTTAGTTGGTTAGAATACTTTCTTCATAGTCACATT
CTCTCAACCTATAATTTGGAATATTGTTGTGGTCTTTTGTCTTTCTCTTAGTATAGCATTTTAA
AAAAATATAAAGCTACCAATCTTTGTACAATTTGTAAATGTTAAGAATTTTTTTTATATCTGT
TAAATAAAAATTATTTCCAACA

FIGURE 122

MRPQGPAASPQRLRGLLLLLLLQLPAPSSASEIPKGGKQKAQLRQREVVDLYNGMCLQGPAQVPGR
DGSPGANVIPGTPGIPGRDGFKEGKECLRESFEESWTPNYKQCSWSSLNYGIDLGKIAECTFTK
MRSNSALRVLFSGSLRLKCRNACCQRWYFTFNGAECSGPLPIEAIYLDQGSPEMNSTINIHRTS
SVEGLCEGIGAGLVDVAIWVGTCSDYPKGDASTGWNSVSRIIEELPK

Signal peptide:

amino acids 1-30

Transmembrane domain:

amino acids 195-217

FIGURE 123

GCTGAGCGTGTGCGCGGTACGGGGCTCTCCTGCCTTCTGGGCTCCAACGCAGCTCTGTGGCTGAA
CTGGGTGCTCATCACGGGAACGTCTGGGCTATGGAATACAGATGTGGCAGCTCAGGTAGCCCCAA
ATTGCCTGGAAGAATACATCATGTTTTTCGATAAGAAGAAATTGTAGGATCCAGTTTTTTTTTA
ACCGCCCCCTCCCCACCCCCCAAAAAAAGCTGTAAAGATGCAAAAACGTAATATCCATGAAGATCC
TATTACCTAGGAAGATTTTGATGTTTTGCTGCGAATGCGGTGTTGGGATTTATTTGTTCTTGGAG
TGTTCTGCGTGGCTGGCAAAGAATAATGTTCCAAAATCGGTCCATCTCCAAGGGGTCCAATTTT
TCTTCTGGGTGTGAGCGAGCCCTGACTCACTACAGTGCAGCTGACAGGGGCTGTGATGCAACTG
GCCCCAAGCCAAAGCAAAAGACCTAAGGACGACCTTTGAACAATACAAAGGATGGGTTTCAATG
TAATTAGGCTACTGAGCGGATCAGCTGTAGCACTGGTTATAGCCCCCACTGTCTTACTGACAATG
CTTTCTTCTGCCGAACGAGGATGCCCTAAGGGCTGTAGGTGTGAAGGCAAAATGGTATATTGTGA
ATCTCAGAAATTACAGGAGATACCCTCAAGTATATCTGCTGGTTGCTTAGGTTTGTCCCTTCGCT
ATAACAGCCTTCAAAAACCTTAAGTATAATCAATTTAAAGGGCTCAACCAGCTCACCTGGCTATAC
CTTGACCATAACCATATCAGCAATATTGACGAAAATGCTTTTAAATGGAATACGCAGACTCAAAGA
GCTGATTCTTAGTTCCAATAGAATCTCCTATTTCTTAAACAATACCTTCAGACCTGTGACAAATT
TACGGAACCTTGGATCTGTCTATAATCAGCTGCATTCTCTGGGATCTGAACAGTTTCGGGGCTTG
CGGAAGCTGCTGAGTTTACATTTACGGTCTAACTCCCTGAGAACCATCCCTGTGCGAATATTCCA
AGACTGCCGCAACCTGGAACCTTTGGACCTGGGATATAACCGGATCCGAAGTTTAGCCAGGAATG
TCTTTGCTGGCATGATCAGACTCAAAGAACCTTCACCTGGAGCACAATCAATTTTCCAAGCTCAAC
CTGGCCCTTTTCCAAGGTTGGTCAAGCCTTCAGAACCTTTACTTGCAGTGAATAAAATCAGTGT
CATAGGACAGACCATGTCTGGACCTGGAGCTCCTTACAAAGGCTTGATTATCAGGCAATGAGA
TCGAAGCTTTTCAAGTGGACCCAGTGTTTTCCAGTGTTGCCGAATCTGCAGCGCCTCAACCTGGAT
TCCAACAAGCTCACATTTATTGGTCAAGAGATTTGGATTCTTGGATATCCCTCAATGACATCAG
TCTTGCTGGGAATATATGGGAATGCAGCAGAAATATTGCTCCCTTGTAAGTGGCTGAAAAGTT
TTAAAGGTCTAAGGGAGAATACAATTATCTGTGCCAGTCCCAAAGAGCTGCAAGGAGTAAATGTG
ATCGATGCAGTGAAGAACTACAGCATCTGTGGCAAAAGTACTACAGAGAGGTTTGATCTGGCCAG
GGCTCTCCCAAAGCCGACGTTTAAAGCCCAAGCTCCCCAGGCCGAAGCATGAGAGCAAACCCCTT
TGCCCCCGACGGTGGGAGCCACAGAGCCCGCCAGAGACCGATGCTGACGCCGAGCACATCTCT
TTCCATAAAATCATCGCGGGCAGCGTGGCGCTTTTCTGTCCGTGCTCGTCATCTGCTGGTTAT
CTACGTGTGATGGAAGCGGTACCCTGCGAGCATGAAGCAGCTGCAGCAGCGCTCCCTCATGCGAA
GGCACAGGAAAAAGAAAAGACAGTCCCTAAAGCAAATGACTCCAGCACCCAGGAATTTTATGTA
GATTATAAACCACCAACACGGAGACCAGCGAGATGCTGCTGAATGGGACGGGACCCTGCACCTA
TAACAAATCGGGCTCCAGGGAGTGTGAGGTATGCAACCATTGTGATAAAAAGAGCTCTTAAAGCT
GGGAAATAAGTGGTCTTTATTGAACTCTGGTGACTATCAAGGGAACGCGATGCCCCCCTCCCC
TTCCCTCTCCCTCTCACTTTGGTGGAAGATCCTTCTTGTCCGTTTTAGTGCATTTCATAATACT
GGTCATTTTCTCTCATACATAATCAACCCATTGAAATTTAAATACCACAATCAATGTGAAGCTT
GAACCTCCGGTTTAAATATAATACCTATTGTATAAGACCCTTTACTGATTCCATTAAATGTCGATTT
GTTTTAAGATAAAACTTCTTTTCATAGGTAAAAA

FIGURE 124

MGFNVIRLLSGSAVALVIAPTPLLTLSSAERGCPKGCRCGKVMYCESQKLQEIPSSISAGCLG
LSLRYNSLQKLKYNQFKGLNQLTWLYLDHNNHISNIDENAFNGIRRLKELILSSNRISYFLNNTFR
PVTNLRNLDLSYNQLHSLGSEQFRGLRKLLSLHLRSNSLRTIPVRI FQDCRNLELLDLGYNRIRS
LARNVFAGMIRLKLHLEHNQFSKLNLFPRVLVSLQNLVYLQWNKISVIGQTMSTWSSLQRLDL
SGNEIEAFSGPSVFQCVPNLQRLNLDNKLTFIGQEILDSWISLNDISLAGNIWECSRNICSLVN
WLKSEFKGLRENTIICASPKELQGVNVIDAVKNYSICGKSTTERFDLARALPKPTFKPKLPRPKHE
SKPPLPPTVGATEPGPETDADAEHISFHKIIAGSVALFLSVLVILLVIYVSWKRYPASMKQLQQR
SLMRRHRKKRQSLKQMTPTSTQEFYVDYKPTNTETSEMLLNGTGPCCTYNKSGSRECEV

Important features of the protein:

Signal peptide:

amino acids 1-33

Transmembrane domain:

amino acids 420-442

N-glycosylation sites.

amino acids 126-129, 357-360, 496-499, 504-507

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 465-468

Tyrosine kinase phosphorylation site.

amino acids 136-142

N-myristoylation sites.

amino acids 11-16, 33-38, 245-250, 332-337, 497-502, 507-512

FIGURE 125

CCGTTATCGTCTTGCGCTACTGCTGAATGTCCGTCCCGGAGGAGGAGGAGAGGCTTTTGCCGCTG
ACCCAGAGATGGCCCCGAGCGAGCAAATTCCTACTGTCCGGCTGCGCGGCTACCGTGGCCGAGCT
AGCAACCTTTCCCTGGATCTCACAAAACTCGACTCCAAATGCAAGGAGAAGCAGCTCTTGCTC
GGTTGGGAGACGGTGCAAGAGAATCTGCCCCCTATAGGGGAATGGTGCGCACAGCCCTAGGGATC
ATTGAAGAGGAAGGCTTTCTAAAGCTTTGGCAAGGAGTGACACCCGCCATTTACAGACACGTAGT
GTATTCTGGAGGTCGAATGGTCACATATGAACATCTCCGAGAGGTTGTGTTGGCAAAAGTGAAG
ATGAGCATTATCCCCTTTGGAAATCAGTCATTGGAGGGATGATGGCTGGTGTATTGGCCAGTTT
TTAGCCAATCCAACCTGACCTAGTGAAGGTTGAGATGCAAAATGGAAGGAAAAAGGAAACTGGAAGG
AAAACCATTGCGATTTTCGTGGTGTACATCATGCATTTGCAAAAATCTTAGCTGAAGGAGGAATAC
GAGGGCTTTGGGCAGGCTGGGTACCCAATATACAAAGAGCAGCACTGGTGAATATGGGAGATTTA
ACCACTTATGATACAGTGAAACACTACTTGGTATTGAATACACCACTTGAGGACAATATCATGAC
TCACGGTTTATCAAGTTTATGTTCTGGACTGGTAGCTTCTATTCTGGGAACACCAGCCGATGTCA
TCAAAAGCAGAATAATGAATCAACCACGAGATAAACAAGGAAGGGGACTTTTGTATAAATCATCG
ACTGACTGCTTGATTCAGGCTGTTCAAGGTGAAGGATTCATGAGTCTATATAAAGGCTTTTACC
ATCTTGGCTGAGAATGACCCCTTGGTCAATGGTGTCTGGCTTACTTATGAAAAATCAGAGAGA
TGAGTGGAGTCAGTCCATTTTAA

FIGURE 126

MSVPEEEERLLPLTQRWPRASKFLLSGCAATVAELATFPLDLTKTRLQMGEAALARLGDGARES
APYRCMVRTALGIEEEGFLKLWQGVTPAIYRHVVYSGGRMVTYEHLEVVFGKSEDEHYPLWKS
VIGGMMAGVIGQFLANPTDLVKVQMMEGKRKLEGKPLRFRGVHHAFAKILAEGGIRGLWAGWVP
NIQRAALVNMGDLTTYDTVKHYLVLNTPLEDNIMTHGLSSLCSGLVASILGTPADVIKSRIMNQP
RDKQGRGLLYKSSTDCLIQAVQGEGFMSLYKGFLPSWLRMTPWSMVFWLTYEKIREMSGVSPF

Transmembrane domains:

amino acids 25-38, 130-147, 233-248

FIGURE 127

CGCGGATCGGACCCAAGCAGGTCGGCGGGCGGCGGACAGGAGAGCGGCCGGGCGTCAGCTCCTCGAC
CCCCGTGTCGGGCTAGTCCAGCGAGGCGGACGGGCGGCGTGGGCCCATGGCCAGGCCCGGCATGG
AGCGGTGGCGCGACCGGCTGGCGCTGGTGACGGGGGCCTCGGGGGGCATCGGCGCGGCCGTGGCC
CGGGCCCTGGTCCAGCAGGGACTGAAGGTGGTGGGCTGCGCCCGCACTGTGGGCAACATCGAGGA
GCTGGCTGCTGAATGTAAGAGTGCAGGCTACCCCGGGACTTTGATCCCCCTACAGATGTGACCTAT
CAAATGAAGAGGACATCCTCTCCATGTTCTCAGCTATCCGTTCTCAGCACAGCGGTGTAGACATC
TGCATCAACAATGCTGGCTTGGCCCCGGCCTGACACCCTGCTCTCAGGCAGCACCAGTGGTTGGAA
GGACATGTTCAATGTGAACGTGCTGGCCCTCAGCATCTGCACACGGGAAGCCTACCAGTCCATGA
AGGAGCGGAATGTGGACGATGGGCACATCATTAACATCAATAGCATGTCTGGCCACCAGTGTTA
CCCCCTGTCTGTGACCCACTTCTATAGTGCCACCAAGTATGCCGTCACTGCGCTGACAGAGGGACT
GAGGCAAGAGCTTCGGGAGGGCCAGACCCACATCCGAGCCACGTGCATCTCTCAGGTGTGGTGG
AGACACAATTGCGCTTCAAACCTCCACGACAAGGACCCTGAGAAGGCAGCTGCCACCTATGAGCAA
ATGAAGTGTCTCAAACCGAGGATGTGGCCGAGGCTGTTATCTACGTCCTCAGCACCCCCGCACA
CATCCAGATTGGAGACATCCAGATGAGGCCCCAGGAGCAGGTGACCTTAGTGACTGTGGGAGCTCC
TCCTTCCCTCCCCACCCTTCATGGCTTGCCTCCTGCCTCTGGATTTTAGGTGTTGATTCTGGAT
CACGGGATACCATTCTGTCCACACCCGACCAGGGGCTAGAAAATTTGTTTGGATTTTATA
TCATCTTGTCAAATTGCTTCAGTTGTAAATGTGAAAAATGGGCTGGGGAAAGGAGTGGTGTCCC
TAATTGTTTTACTTGTTAACTTGTCTTGTGCCCTGGGCACTTGGCCTTGTCTGCTCTCAGTG
TCTTCCCTTTGACATGGGAAAGGAGTTGTGGCCAAAATCCCCATCTTCTTGACCTCAACGTCTG
TGGCTCAGGGCTGGGGTGGCAGAGGGAGGCCTTACCTTATATCTGTGTTGTTATCCAGGGCTCC
AGACTTCCTCCTCTGCCTGCCCCACTGCACCCTCTCCCCCTTATCTATCTCCTTCTCGGCTCCCC
AGCCCAGTCTTGGCTTCTTGTCCCCTCCTGGGGTCATCCCTCCACTCTGACTCTGACTATGGCAG
CAGAACACCAGGGCCTGGCCCAGTGGATTTCATGGTGATCATTAAAAAGAAAAATCGCAACCAA
AAAAAAAAA

FIGURE 128

MARPGMERWRDRDLALVTGASGGIGA AAVARALVQQGLKVVG CARTVGNIEELAAECKSAGYPGTLI
PYRCDLSNEEDILSMFSAIRSQHSGVDICINNAGLARPD TLLSGSTSGWKDMFNVNVLALSICTR
EAYQSMKERNVDDGHIININMSGHRVLP LSVTHFY SATKYAVTALTEGLRQELREAQT HIRATC
ISPGVVETQFAFKLHDKDPEKAAATYEQMKCLKPEDVAEAVIYVLSTPAHIQIGDIQMRPTEQVT

Important features of the protein:

Signal peptide:

amino acids 1-17

N-myristoylation sites.

amino acids 18-24, 21-27, 22-28, 24-30, 40-46, 90-96, 109-115,
199-205

Short-chain alcohol dehydrogenase.

amino acids 30-42, 104-114

FIGURE 129

AACTTCTACATGGGCCTCCTGCTGCTGGTGCTCTTCCTCAGCCTCCTGCCGGTGGCCTACACCAT
CATGTCCCTCCCACCTCCTTTGACTGCGGGCCGTTTCAGGTGCAGAGTCTCAGTTGCCCGGGAGC
ACCTCCCTCCCAGGCAGTCTGCTCAGAGGGCCTCGGCCAGAATTCAGTTCTGGTTTCATGC
CAGCCTGTAAAAGGCCATGGAACCTTTGGGTGAATCACCGATGCCATTTAAGAGGGTTTTCTGCCA
GGATGGAAATGTTAGGTCGTTCTGTGTCTGCGCTGTTCAATTCAGTAGCCACCAGCCACCTGTGG
CCGTTGAGTGCTTGAAATTGAGGAACTGAGAAAATTAATTTCTCATGTATTTTTCTCATTTATTTA
TTAATTTTTAACTGATAGTTGTACATATTTGGGGGTACATGTGATATTTGGATACATGTATACAA
TATATAATGATCAAAATCAGGGTAACTGGGATATCCATCACATCAAACATTTATTTTTTATTCTTT
TTAGACAGAGTCTCACTCTGTCAACCAGGCTGGAGTGCAGTGGTGCCATCTCAGCTTACTGCAAC
CTCTGCCTGCCAGGTTCAAGCGATTCTCATGCCTCCACCTCCCAAGTAGCTGGGACTACAGGCAT
GCACCACAATGCCCACTAATTTTTGTATTTTAGTAGAGACGGGGTTTTGCCATGTTGCCCAGG
CTGGCCTTGAACCTCCTGGCCTCAAACAATCCACTTGCCTCGGCCTCCCAAAGTGTTATGATTACA
GGCGTGAGCCACCGTGCCTGGCCTAAACATTTATCTTTCTTTGTGTTGGGAACTTTGAAATTAT
ACAATGAATTATTGTTAACTGTCATCTCCCTGCTGTGCTATGGAACACTGGGACTTCTTCCCTCT
ATCTAACTGTATATTTGTACCAGTTAACCAACCGTACTTCATCCCCACTCCTCTCTATCCTTCCC
AACCTCTGATCACCTCATTTCTACTCTCTACCTCCATGAGATCCACTTTTTTAGCTCCCATGTG
AGTAAGAAAATGCAATATTTGTCTTTCTGTGCCTGGCTTATTTCACTTAACATAATGACTTCCTG
TTCCATCCATGTTGCTGCAATGACAGGATTCGTTCTTAATTTCAATTAAAATAACCACACATG
GCAAAAA

FIGURE 130

MGLLLLVLFLSLLPVAYTIMSLPPSFDCGPFCRVSVAREHLPSRGSLLRGPRPRIPVLVSCQPV
KGHGTLGESMPFKRVCQDGNVRSFCVCAVHFSSHQPPVAVECLK

Important features of the protein:

Signal peptide:

amino acids 1-18

N-myristoylation site.

amino acids 86-92

Zinc carboxypeptidases, zinc-binding region 2 signature.

amino acids 68-79

FIGURE 131

TTCTGAAGTAACGGAAGCTACCTTGTATAAAGACCTCAACACTGCTGACCATGATCAGCGCAGCCTGGAGC
 ATCTTCCTCATCGGGACTAAAATTGGGCTGTTCCCTTCAAGTAGCACCTCTATCAGTTATGGCTAAATCCTG
 TCCATCTGTGTGTCGCTGCGATGCGGGTTTCATTTACTGTAATGATCGCTTTCTGACATCCATTCCAACAG
 GAATACCAGAGGATGCTACAACCTCTCTACCTTCAGAACAACCAAATAAATAATGCTGGGATTCCTTCAGAT
 TTGAAAACTTGCTGAAAGTAGAAGAATATACCTATACCACAACAGTTTAGATGAATTTCTACCAACCT
 CCCAAGTATGTAAAAGAGTTACATTTGCAAGAAAATAACATAAGGACTATCACTTATGATTCACTTTCAA
 AAATTCCTATCTGGAAGAATTACATTTAGATGACAACCTCTGTCTCTGCAGTTAGCATAGAAGAGGGAGCA
 TTCCGAGACAGCAACTATCTCCGACTGCTTTTCTGTCCCGTAATCACCTTAGCACAATTCCTCGGGTTT
 GCCCAGGACTATAGAAGAACTACGCTTGGATGATAATCGCATATCCACTATTTTCATCACCATCTCTTCAAG
 GTCTCACTAGTCTAAAACGCTGTTTCTAGATGGAACCTGTTGAACAATCATGGTTTAGGTGACAAAGTT
 TTCTTCAACCTAGTTAATTGACAGAGCTGTCCCTGGTGCAGGAAATTCCTGACTGCTGCACCAGTAAACCT
 TCCAGGCACAACTGAGGAAGCTTTATCTTCAAGATAACCACATCAATCGGGTCCCCCAATGCTTTTT
 CTATCTAAGGCAGCTCTATCGACTGGATATGTCCAATAATAACCTAAGTAATTTACCTCAGGGTATCTTT
 GATGATTTGGACAATATAACACAACCTGATTCTTCGCAACAATCCCTGGTATTGCGGGTGCAAGATGAAATG
 GGTACGTGACTGGTTACAATCACTACCTGTGAAGGTCAACGTGCGTGGGCTCATGTGCCAAGCCCCAGAAA
 AGGTTCTGTTGGGATGGCTATTAAAGGATCTCAATGCAGAACTGTTTGATTGTAAGGACAGTGGGATTGTAAGC
 ACCATTGAGATAACCACTGCAATACCCAACACAGTGTATCCTGCCCAAGGACAGTGGCCAGCTCCAGTGAC
 CAAACAGCCAGATATTAAGAACCCCAAGCTCACTAAGGATCAACAAACCACAGGGAGTCCCTCAAGAAAAA
 CAATTACAATTACTGTGAAGTCTGTACCTCTGATACCATTCATATCTCTTGGAACCTTGCTCTACCTATG
 ACTGCTTTGAGACTCAGCTGGCTTAAACTGGGCCATAGCCCGGCATTTGGATCTATAACAGAAACAATTGT
 AACAGGGGAACGCAGTGAGTACTTGGTCACAGCCCTGGAGCCTGATTCACCTATAAAGTATGCATGGTTC
 CCATGGAACCCAGCAACCTCTACCTATTTGATGAAACTCCTGTTTGTATTGAGACTGAAACTGCACCCCTT
 CGAATGTACAACCCTACAACCACCTCAATCGAGAGCAAGAGAAAGAACCTTACAAAACCCCAATTTACC
 TTTGGCTGCCATCATTTGGTGGGGCTGTGGCCCTGGTTACCATTGCCCTTCTTGCTTTAGTGTGTGGTATG
 TTCATAGGAATGGATCGCTCTTCTCAAGGAAGTGTGCATATAGCAAAGGGAGGAGAAGAAAGGATGACTAT
 GCAGAAGCTGGCACTAAGAAGGACAACCTCTATCCTGGAATCAGGGAACTTCTTTTCAGATGTTACCAAT
 AAGCAATGAACCCATCTCGAAGGAGGAGTTTGTAAATACACACCATATTTCTCCTAATGGAATGAATCTGT
 ACAAAAACAATCACAGTGAAAGCAGTAGTAACCGAAGCTACAGAGACAGTGGTATTCAGACTCAGATCAC
 TCACACTCATGATGCTGAAGGACTCACAGCAGACTTGTGTTTTGGGTTTTTTAAACCTAAGGGAGGTGATG
 GT

FIGURE 132

MISAAWSIFLIGTKIGLFLQVAPLSVMAKSCPSVCRCDAGFIYCNDRFLTSIPTGIPEDATTLYL
QNNQINNAGIPSDLKNLLKVERIYLYHNSLDEFPTNLPKYVKELHLQENNIRTITYDSLSKIPYL
EELHLDNSVSAVSIEEGAFRDSNYLRLLFLSRNHLSTIPWGLPRTIEELRLDDNRISTISSPSL
QGLTSLKRLVLDGNLLNNHGLGDKVFFNLVNLTELSLVRNSLTAAPVNLPGTNLRKLYLQDNHIN
RVPPNAFSYLRQLYRLDMSNNNLSNLPQGI FDDL DNITQLILRNNPWYCGCKMKWVRDWLQSLPV
KVNVRGLMCQAPEKVRGMAIKDLNAELFDCCKDSGIVSTIQITTAIPNTVYPAQQQWPAPVTKQPD
IKNPKLTKDQQTGSPSRKTTITITVKSVTSDTIHISWKLALPMTALRLSWLKLGHSPAFGSITET
IVTGERSEYLVTALEPDSPYKVCMPMETSNNLYLFDETPVCIETETAPLRMYNPPTTLNREQEKE
PYKNPNPLAAIIGGAVALVTIALALVCWYVHRNGSLFSRNCAYSKGRRRKDDYAEAGTKKDNS
ILEIRETSFQMLPISNEPISKEEFVIHTIFPPNGMNLYKNNHSESSSNRSYRDSGIPDSDHSHS

Important features of the protein:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 531-552

N-glycosylation sites.

amino acids 226-229, 282-285, 296-299, 555-558, 626-629, 633-636

Tyrosine kinase phosphorylation site.

amino acids 515-522

N-myristoylation sites.

amino acids 12-17, 172-177, 208-213, 359-364, 534-539, 556-561,
640-645

Amidation site.

amino acids 567-570

Leucine zipper pattern.

amino acids 159-180

Phospholipase A2 aspartic acid active site.

amino acids 34-44

FIGURE 133

CCGTCATCCCCCTGCAGCCACCCTTCCCAGAGTCCTTTGCCCAGGCCACCCAGGCTTCTTGCCA
GCCCTGCCGGGCCACTTGTCTTCATGCTCTGCCAGGGGGAGGTGGGAAGGAGGTGGGAGGAGGGCG
TGCAGAGGCAGTCTGGGCTTGGCCAGAGCTCAGGGTGCTGAGCGTGTGACCAGCAGTGAGCAGAG
GCCGGCCATGGCCAGCCTGGGGCTGCTGCTCCTGCTCTTACTGACAGCACTGCCACCGCTGTGGT
CCTCCTCACTGCCTGGGCTGGACACTGCTGAAAGTAAAGCCACCATTGCAGACCTGATCCTGTCT
GCGTGGAGAGAGCCACCGTCTTCTAGAACAGAGGCTGCCTGAAATCAACCTGGATGGCATGGT
GGGGGTCCGAGTGCTGGAAGAGCAGCTAAAAAGTGTCGGGAGAAGTGGGCCCAGGAGCCCTGC
TGCAGCCGCTGAGCCTGCGCGTGGGGATGCTGGGGGAGAAGCTGGAGGCTGCCATCCAGAGATCC
CTCCACTACCTCAAGCTGAGTGATCCCAAGTACCTAAGAGAGTTCCAGCTGACCTCCAGCCCGG
GTTTTGGAAGCTCCACATGCCTGGATCCACACTGATGCCTCCTTGGTGTACCCACGTTCTGGGC
CCCAGGACTCATTCTCAGAGGAGAGAAGTGACGTGTGCCTGGTGCAGCTGCTGGGAACCGGGACG
GACAGCAGCGAGCCCTGCGGCCTCTCAGACCTCTGCAGGAGCCTCATGACCAAGCCCGGCTGCTC
AGGCTACTGCCTGTCCACCAACTGCTCTTCTCTCTGCGCCAGAATGAGGGGATGCACACAGG
GACCACTCCAACAGAGCCAGGACTATATCAACCTCTTCTGCGCCAACATGATGGACTTGAACCGC
AGAGCTGAGGCCATCGGATACGCCCTACCCTACCCGGGACATCTTCATGGAAAACATCATGTTCTG
TGGAATGGGCGGCTTCTCCGACTTCTACAAGCTCCGGTGGCTGGAGGCCATTCTCAGCTGGCAGA
AACAGCAGGAAGGATGCTTCGGGGAGCCTGATGCTGAAGATGAAGAATTATCTAAAGCTATTCAA
TATCAGCAGCATTTTTCGAGGAGAGTGAAGAGGCGAGAAAAACAATTTCCAGATTCTCGCTCTGT
TGCTCAGGCTGGAGTACAGTGGCGCAATCTCGGCTCACTGCAACCTTTGCCTCCTGGGTTCAGC
AATTCTCTTGCTCATCTCTCCGAGTAGCTGGGACTACAGGAGCGTGCCACCATACTGGCTAAT
TTTTATATTTTTTTAGTAGAGACAGGGTTTCATCATGTTGCTCATGCTGGTCTCGAACTCCTGAT
CTCAAGAGATCCGCCACCTCAGGCTCCCAAAGTGTTGGGATTATAGGTGTGAGCCACCGTGTCTG
GCTGAAAAGCACTTTCAAAGAGACTGTGTTGAATAAAGGGCCAAGGTTCTTGCCACCCAGCACTC
ATGGGGGCTCTCTCCCTAGATGGCTGCTCCTCCCAACACAGCCACAGCAGTGGCAGCCCTGG
GTGGCTTCCTATACATCCTGGCAGAATACCCCCAGCAAACAGAGGCCACACCCATCCACACCG
CCACCACCAAGCAGCCGCTGAGACGGACGGTTCCATGCCAGCTGCCTGGAGGAGGAACAGACCCC
TTTAGTCCTCATCCCTTAGATCCTGGAGGGCACGGATCACATCCTGGGAAGAAGGCATCTGGAGG
ATAAGCAAAGCCACCCCGACACCCAATCTTGAAGCCCTGAGTAGGCAGGGCCAGGGTAGGTGGG
GGCCGGGAGGGACCCAGGTGTGAACGGATGAATAAAGTTCAACTGCAACTGAAAAAAAAAA

FIGURE 134

MSARGRWEGGRRACRGS LGLARAQGAERVTSSSEQRPAMASLGLLLLLLLLTALPPLWSSSLPGLD
TAESKATIADLILSALERATVFLEQRLPEINLDGMVGVRVLEEQLKSVREKWAQEPLLQPLSLRV
GMLGEKLEAAIQRS LHYLKLSDPKYLREFQLTLQPGFWKLP HAWIHTDASLVYPTFGPQDSFSEE
RSDVCLVQLLGTGDSSEPCGLSDLCRSLMTKPGCSGYCLSHQLLFFLWARMRGCTQGPLQQSQD
YINLFCANMMDLNRRAEAIGYAYPTRDIFMENIMFCGMGGFSDFYKLRWLEAILSWQKQEGCFG
EPDAEDEELSKAIQYQQHFSRRVKRREKQFPDSRSVAQAGVQWRNLGSLQPLPPGFKQFSLILP
SSWDYRSVPPYLANFYIFLVETGFHHVAHAGLELLISRPPTSGSQSVGL

Important features of the protein:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 39-56

Tyrosine kinase phosphorylation sites.

amino acids 149-156, 274-282

N-myristoylation sites.

amino acids 10-16, 20-26, 63-69, 208-214

Amidation site.

amino acids 10-14

Glycoprotein hormones beta chain signature 1.

amino acids 230-237

FIGURE 135

GGTCTGAGTGCAGAGCTGCTGTCATGGCGGCCGCTCTGTGGGGCTTCTTTCCCGTCCTGCTGCTG
CTGCTGCTATCGGGGGATGTCCAGAGCTCGGAGGTGCCCGGGGCTGCTGCTGAGGGATCGGGAGG
GAGTGGGGTCGGCATAGGAGATCGCTTCAAGATTGAGGGGCGTGAGTTGTTCCAGGGGTGAAGC
CTCAGGACTGGATCTCGGCGGCCCGAGTGTGGTAGACGGAGAAGAGCACGTGGTTTCTTTAAG
ACAGATGGGAGTTTTGTGGTTCATGATATACCTTCTGGATCTTATGTAGTGGAAGTTGTATCTCC
AGCTTACAGATTTGATCCCGTTCGAGTGGATATCACTTCGAAAGGAAAAATGAGAGCAAGATATG
TGAATTACATCAAAACATCAGAGGTTGTGAGACTGCCCTATCCTCTCCAAATGAAATCTTCAGGT
CCACCTTCTTACTTTATTAAAAAGGAATCGTGGGGCTGGACAGACTTTCTAATGAACCCAATGGT
TATGATGATGGTTCTTCCTTTATTGATATTTGTGCTTCTGCCTAAAGTGGTCAACACAAGTGATC
CTGACATGAGACGGGAAATGGAGCAGTCAATGAATATGCTGAATTCCAACCATGAGTTGCCTGAT
GTTTCTGAGTTCATGACAAGACTCTTCTCTTCAAAATCATCTGGCAAATCTAGCAGCGGCAGCAG
TAAACAGGCAAAAGTGGGGCTGGCAAAGGAGGTAGTCAGGCCGTCCAGAGCTGGCATTTCAC
AAACACGGCAACACTGGGTGGCATCCAAGTCTTGAAAACCGTGTGAAGCAACTACTATAAACTT
GAGTCATCCCGACGTTGATCTCTTACAACGTGTGTATGTT
AACTTTTTAGCACATGTTTTGTACTTGGTACACGAGAAAACCCAGCTTTCATCTTTTGTCTGTAT
GAGGTCAATATTGATGTCACCTGAATTAATTACAGTGTCCCTATAGAAAATGCCATTAATAAATTAT
ATGAACTACTATACATTATGTATATTAATTAACATCTTAATCCAGAAATCAAAAAAAAAAAAA
AAAAAAAAAAAAAAAA

FIGURE 136

MAAALWGFFPVLLLLLLSGDVQSSEVPGAAAEGLSGSGVGIGDRFKIEGRAVVPGVKPDWISAA
RVLVDGEEHVGFLKTDGSFVVHDIPSGSYVVEVVSPAYRFDPPVRVDITSKGKMRARYVNYIKTSE
VVRLPYPLQMKSSGPPSYFIKRESWGWTDFLMNPMVMMVLPLLIFVLLPKVVNTSDPDMRREME
QSMNMLNSNHELDPDVSEFMTRLFSSKSSGKSSSGSSKTGKSGAGKRR

Important features of the protein:

Signal sequence:

amino acids 1-23

Transmembrane domain:

amino acids 161-182

N-glycosylation site.

amino acids 184-187

Glycosaminoglycan attachment sites.

amino acids 37-40, 236-239

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 151-154

N-myristoylation sites.

amino acids 33-38, 36-41, 38-44, 229-234

Amidation site.

amino acids 238-241

ATP/GTP-binding site motif A (P-loop).

amino acids 229-236

FIGURE 137

GATGGCGCAGCCACAGCTTCTGTGAGATTCGATTTCTCCCCAGTTCCTGTGGGTCTGAGGGGA
CCAGAAGGGTGAGCTACGTTGGCTTTCTGGAAGGGGAGGCTATATGCGTCAATCCCCAAACAA
GTTTTGACATTTCCCCTGAAATGTCATTCTCTATCTATTCACTGCAAGTGCCTGCTGTTCCAGGC
CTTACCTGCTGGGCACTAACGGCGGAGCCAGGATGGGGACAGAATAAAGGAGCCACGACCTGTGC
CACCAACTCGCACTCAGACTCTGAACTCAGACCTGAAATCTTCTCTTACGGGAGGCTTGGCAGT
TTTTCTTACTCCTGTGGTCTCCAGATTTCAGGCCTAAGATGAAAGCCTCTAGTCTTGCCTTCAGC
CTTCTCTCTGCTGCGTTTTATCTCCTATGGACTCCTTCCACTGGACTGAAGACACTCAATTTGGG
AAGCTGTGTGATCGCCACAAACCTTCAGGAAATACGAAATGGATTTTCTGAGATACGGGGCAGTG
TGCAAGCCAAAGATGGAAACATTGACATCAGAATCTTAAGGAGGACTGAGTCTTTGCAAGACACA
AAGCCTGCGAATCGATGCTGCCTCCTGCGCCATTTGCTAAGACTCTATCTGGACAGGGTATTTAA
AACTACCAGACCCCTGACCATTATACTCTCCGGAAGATCAGCAGCCTCGCCAATTCTTTCTTA
CCATCAAGAAGGACCTCCGGCTCTCTCATGCCACATGACATGCCATTGTGGGGAGGAAGCAATG
AAGAAATACAGCCAGATTCTGAGTCACTTTGAAAAGCTGGAACCTCAGGCAGCAGTTGTGAAGGC
TTTGGGGGAAC TAGACATTCTTCTGCAATGGATGGAGGAGACAGAA TAGGAGGAAAGTGATGCTG
CTGCTAAGAATATTGAGGTCAAGAGCTCCAGTCTTCAATACCTGCAGAGGAGGCATGACCCCAA
ACCACCATCTCTTTACTGTACTAGTCTTGTGCTGGTCACAGTGTATCTTATTTATGCATTACTTG
CTTCCTTGCAATGATTGTCTTTATGCATCCCCAATCTTAATTGAGACCATACTTGTATAAGATTTT
TGTAATATCTTTCTGCTATTGGATATATTTATTAGTTAATATATTTATTTATTTTTTGCTATTTA
ATGTATTTATTTTTTTACTTGGACATGAACTTTAAAAAAATTCACAGATTATATTTATAACCTG
ACTAGAGCAGGTGATGTATTTTTATACAGTAAAAAAAACCTTGTAATTC TAGAAGAGTGG
CTAGGGGGGTTATTCATTTGTATTCAACTAAGGACATATTTACTCATGCTGATGCTCTGTGAGAT
ATTTGAAATTGAACCAATGACTACTTAGGATGGGTTGTGGAATAAGTTTGTATGTGGAATTGCAC
ATCTACCTTACAATTACTGACCATCCCCAGTAGACTCCCCAGTCCCATAATTGTGTATCTTCCAG
CCAGGAATCCTACACGGCCAGCATGTATTTCTACAAATAAAGTTTCTTTGCATACCAAAAAAAA
AAAAAAAAAA

FIGURE 138

MRQFPKTSFDISPEMSFSIYSLQVPAVPGLTCWALTAEPGWGQNGATTTCATNSHSDSELRPEIF
SSREAWQFFLLWSPDFRPKMKASSLAFSLLSAFYLLWTPSTGLKTLNLGSCVIATNLQEIRNG
FSEIRGSVQAKDGNIDIRILRRTESLQDTKPANRCCLLRHLLRLYLDRVFKNYQTPDHYTLRKIS
SLANSFLTIKDLRLSHAHTCHCGEEAMKKYSQILSHFEKLEPQAAVVKALGELDILLQWMEET
E

Important features of the protein:

Signal peptide:

amino acids 1-42

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 192-195, 225-228

N-myristoylation sites.

amino acids 42-47, 46-51, 136-141

FIGURE 139

CCTGGAGCCGGAAGCGCGGCTGCAGCAGGCGGAGGCTCCAGGTGGGGTCGGTTCGCATCCAGCC
TAGCGTGTCCACGATCGCGGCTGGGCTCCGGGACTTTCGCTACCTGTTGCGTAGCGATCGAGGTGC
TAGGGATCGCGGTCTTCTTCGGGGATTCTTCCCGGCTCCCGTTCGTTCCTCTGCCAGAGCGGAA
CACGGAGCGGAGCCCCAGCGCCCGAACCTCGGCTGGAGCCAGTTCTAACTGGACCACGCTGCC
ACCACCTCTCTTCAGTAAAGTTGTTATTGTTCTGATAGATGCCTTGAGAGATGATTTTGTGTTTG
GGTCAAAGGGTGTGAAATTTATGCCCTACACAACCTTACCTTGTGGAAAAAGGAGCATCTCACAGT
TTTGTGGCTGAAGCAAAGCCACCTACAGTTACTATGCCCTGAATCAAGGCATTGATGACGGGGAG
CCTTCCTGGCTTTGTGACGTCATCAGGAACCTCAATTCTCCTGCACTGCTGGAAGACAGTGTGA
TAAGACAAGCAAAAGCAGCTGGAAAAAGAATAGTCTTTTATGGAGATGAAACCTGGGTAAATTA
TTCCCAAAGCATTGTGGAATATGATGGAACAACCTCATTTTCGTGTCAGATTACACAGAGGT
GGATAATAATGTCACGAGGCATTTGGATAAAGTATTAAAAAGAGGAGATTGGGACATATTAATCC
TCCACTACCTGGGCTGGACCACATTGGCCACATTTAGGGCCCAACAGCCCCCTGATTGGGCAG
AAGCTGAGCGAGATGGACAGCGTGCTGATGAAGATCCACACCTCACTGCAGTCGAAGGAGAGAGA
GACGCTTTTACCAATTTGCTGGTTCTTTGTTGGTGACCATGGCATGTCTGAAACAGGAAGTCACG
GGGCTCCTCCACCGAGGAGGTGAATACACCTCTGATTTTAAATCAGTTCTGCGTTTGAAAGGAAA
CCCGGTGATATCCGACATCCAAAGCACGTCCAATAGACGGATGTGGCTGCGACACTGGCGATAGC
ACTTGGCTTACCGATTCCAAAAGACAGTGTAGGGAGCCTCCTATTCCAGTTGTGGGAAGGAAGAC
CAATGAGAGAGCAGTTGAGATTTTACATTTGAATACAGTGCAGCTTAGTAAACTGTTGCAAGAG
AATGTGCCGTGATATGAAAAAGATCCTGGGTTTGAGCAGTTTAAATGTGAGAAAGATTGCATGG
GAACTGGATCAGACTGTACTTGGAGGAAAAGCATTCAGAAAGTCTATTCAACCTGGGCTCCAAGG
TTCTCAGGCAGTACCTGGATGCTCTGAAGACGCTGAGCTTGTCCCTGAGTGCAAAAGTGGCCAG
TTCTCACCTGCTCCTGCTCAGCGTCCCACAGGCAGTGCACAGAAAGGCTGAGCTGGAAGTCCCA
CTGTCATCTCCTGGGTTTTCTCTGCTCTTTTATTTGGTGATCCTGGTTCTTTTCGGCCGTTACGT
CATTTGTGTGCACCTCAGCTGAAAGTTTCGTGCTACTTCTGTGGCCTCTCGTGGCTGGCGGCAGGCT
GCCTTTTCGTTTACCAGACTCTGGTTGAACACCTGGTGTGTGCCAAGTGTGGCAGTGCCTTGGAC
AGGGGGCCTCAGGGAAGGACGTGGAGCAGCCTTATCCAGGCCTCTGGGTGTCCCAGACAGGTG
TTCACATCTGTGCTGTGAGGTGAGATGCCTCAGTTCTTGAAAGCTAGGTTCTCTGCGACTGTTAC
CAAGGTGATTGTAAGAGCTGGCGGTACAGAGGAACAAGCCCCCAGCTGAGGGGGTGTGTGAA
TCGGACAGCCTCCAGCAGAGGTGTGGGAGCTGCAGCTGAGGGAAGAAGAGACAATCGGCCTGGA
CACTCAGGAGGGTCAAAGGAGACTTGGTCGCACCACTCATCTGCCACCCCAAGATGCATCCT
GCCTCATCAGGTCCAGATTTCTTTCCAAGGCGGACGTTTTCTGTTGGAATTCTTAGTCTTGGCC
TCGGACACCTTCAATTCGTTAGCTGGGGAGTGGTGGTGAGGCAGTGAAGAAGAGGCGGATGGTCAC
ACTCAGATCCACAGAGCCAGGATCAAGGGACCCACTGCAGTGGCAGCAGGACTGTTGGGCCCCC
ACCCCAACCCTGCACAGCCCTCATCCCTCTTGGCTTGAGCCGTGAGAGCCCTGTGCTGAGTGT
CTGACCGAGACACTCACAGCTTTGTGATCAGGGCACAGGCTTCCTCGGAGCCAGGATGATCTGTG
CCACGCTTGCACCTCGGGCCCATCTGGGCTCATGCTCTCTCTCTGCTATTGAATTAGTACCTAG
CTGCACACAGTATGTAGTTACCAAAAGAATAAACGGCAATAATTGAGAAAAAAA

FIGURE 140

MRLGSGTFATCCVAIEVLGIAVFLRGFFPAPVRSSARAEGAEPPAPEPSAGASSNWTTLPPPLF
SKVVIVLIDALRDDFVFGSKGVKFMPTTYLVEKGASHSFVAEAKPPTVTMPRIKALMTGSLPGF
VDVIRNLNSPALLEDSEVIRQAKAAGKRIVFYGDETWVKLFPKHFVEYDGTTSFFVSDYTEVDNNV
TRHLDKVLKRGDWDILILHYLGLDGHIGHISGPN SPLIGQKLSEMDSVLMKIHTSLQSKERETPLP
NLLVLCGDHGMSETGSHGASSTEEVNTPLILISSAFERKPGDIRHPKHVQ

Important features of the protein:

Signal peptide:

amino acids 1-34

Transmembrane domain:

amino acids 58-76

N-glycosylation sites.

amino acids 56-60, 194-198

N-myristoylation sites.

amino acids 6-12, 52-58, 100-106, 125-131, 233-239, 270-276,
275-281, 278-284

Amidation site.

amino acids 154-158

Cell attachment sequence.

amino acids 205-208

FIGURE 141

GGCACGAGGCAAGCCTTCCAGGTTATCGTGACGCACCTTGAAAGTCTGAGAGCTACTGCCCTACA
GAAAGTTACTAGTGCCCTAAAGCTGGCGCTGGCACTGATGTTACTGCTGCTGTTGGAGTACAAC
TCCCTATAGAAAACAACTGCCAGCACCTTAAGACCACTCACACCTTCAGAGTGAAGAACTTAAAC
CCGAAGAAATTCAGCATTTCATGACCAGGATCACAAAGTACTGGTCCTGGACTCTGGGAATCTCAT
AGCAGTTCAGATAAAAACCTACATACGCCAGAGATCTTCTTTGCATTAGCCTCATCCTTGAGCT
CAGCCTCTGCGGAGAAAGGAAGTCCGATTCTCCTGGGGGTCTCTAAAGGGGAGTTTGTCTCTAC
TGTGACAAGGATAAAGGACAAAGTCATCCATCCCTTCAGCTGAAGAAGGAGAACTGATGAAGCT
GGCTGCCCCAAAAGGAATCAGCACGCCGGCCCTTCATCTTTTATAGGGCTCAGGTGGGCTCCTGGA
ACATGCTGGAGTCGGCGGCTCACCCCGGATGGTTCATCTGCACCTCCTGCAATTGTAATGAGCCT
GTTGGGGTGACAGATAAATTTGAGAACAGGAAACACATTGAATTTTCATTTCAACCAGTTTGCAA
AGCTGAAATGAGCCCCAGTGAGGTCAGCGATTAGGAAACTGCCCCATTGAACGCCTTCCTCGCTA
ATTTGAACTAATTGTATAAAAAACACCAAACCTGCTCACT

FIGURE 142

MLLLLLLEYNFPIENNCQHLKTTHTFRVKNLNPKKFSIHDQHKVLVLDSGNLIAVPDKNYIRPEI
FFALASSLSSASAEKGSPIILLGVSKGEFCLYCDKDKGQSHPSLQLKKEKLMKLAAQKESARRPFI
FYRAQVGSWNMLESAHPGWFICTSCNCNEPVGVTDFENRKHIEFSFQPVCKAEMSPSEVSD

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 33-36

N-myristoylation site.

amino acids 50-55, 87-92

Interleukin-1

amino acids 37-182

FIGURE 143

CTAGAGAGTATAGGGCAGAAGGATGGCAGATGAGTGACTCCACATCCAGAGCTGCCTCCCTTTAA
TCCAGGATCCTGTCCTTCTGTCTGTAGGAGTGCCTGTTGCCAGTGTGGGGTGAGACAAGTTTG
TCCCACAGGGCTGTCTGAGCAGATAAGATTAAGGGCTGGGTCTGTGCTCAATTAACCTGTGGG
CACGGGGGCTGGGAAGAGCAAAGTCAGCGGTGCCTACAGTCAGCACCATGCTGGGCCTGCCGTGG
AAGGGAGGTCTGTCTGGGCGCTGCTGCTGCTTCTCTTAGGCTCCCAGATCCTGCTGATCTATGC
CTGGCATTTCACGAGCAAAGGGACTGTGATGAACACAATGTCATGGCTCGTTACCTCCCTGCCA
CAGTGGAGTTTGCTGTCCACACATTCAACCAACAGAGCAAGGACTACTATGCCTACAGACTGGGG
CACATCTTGAATTCCTGGAAGGAGCAGGTGGAGTCCAAGACTGTATTCTCAATGGAGCTACTGCT
GGGGAGAACTAGGTGTGGGAAATTTGAAGACGACATTGACAACCTGCCATTTCCAAGAAAGCACAG
AGCTGAACAATACTTTCACCTGCTTCTTCACCATCAGCACCAGGCCCTGGATGACTCAGTTCAGC
CTCCTGAACAAGACCTGCTTGGAGGGATTCCACTTCAGTGAAACCCACTCACAGGCTTGTCATGT
GCTGCTCCACATTCCTGGACATCAGCACTACTCTCCTGAGGACTCTTCAGTGGCTGAGCAGCT
TTGGACTTGTTTGTATCTATTTTGCATGTGTTTGAGATCTCAGATCAGTGTTTTAGAAAATCC
ACACATCTTGAGCCTAATCATGTAGTGTAGATCATTAAACATCAGCATTTTAAGAAAAAAAAA
AA

FIGURE 144

MLGLPWKGGLSWALLLLLLLGSQILLIYAWHFHEQRDCDEHNVMARYLPATVEFAVHTFNQQSKDY
YAYRLGHILNSWKEQVESKTVFSMELLGRTRCGKFEDDIDNCHFQESTELNNTFTCFFTISTRP
WMTQFSLLNKTCLEGFH

Important features of the protein:

Signal peptide:

amino acids 1-25

N-glycosylation sites.

amino acids 117-121, 139-143

N-myristoylation site.

amino acids 9-15

FIGURE 145

CTGTGCAGCTCGAGGCTCCAGAGGCACACTCCAGAGAGAGCCAAGGTTCTGACGCGATGAGGAAG
CACCTGAGCTGGTGGTGGCTGGCCACTGTCTGCATGCTGCTTTCAGCCACCTCTCTGCGGTCCA
GACGAGGGGCATCAAGCACAGAATCAAGTGGAACCGGAAGGCCCTGCCAGCACTGCCAGATCA
CTGAGGCCCAGGTGGCTGAGAACCGCCCGGGAGCCTTCATCAAGCAAGGCCGAAGCTCGACATT
GACTTCGGAGCCGAGGGCAACAGGTACTACGAGGCCAAC TACTGGCAGTTCCCCGATGGCATCCA
CTACAACGGCTGCTCTGAGGCTAATGTGACCAAGGAGGCATTTGTACCGGCTGCATCAATGCCA
CCCAGGCGGCGAACCAGGGGGAGTTCCAGAAGCCAGACAACAAGCTCCACCAGCAGGTGCTCTGG
CGGCTGGTCCAGGAGCTCTGCTCCCTCAAGCATTGCGAGTTTGGTTGGAGAGGGGCGCAGGACT
TCGGGTCACCATGCACCAGCCAGTGCTCCTCTGCCTTCTGGCTTTGATCTGGCTCATGGTGAAAT
AAGCTTGCCAGGAGGCTGGCAGTACAGAGCGCAGCAGCGAGCAAATCCTGGCAAGTGACCCAGCT
CTTCTCCCCCAAACCCACGCGTGTCTGAAGGTGCCAGGAGCGGCGATGCACTCGCACTGCAAA
TGCCGCTCCACGTATGCGCCCTGGTATGTGCCTGCGTTCTGATAGATGGGGGACTGTGGCTTCT
CCGTCACTCCATTCTCAGCCCCCTAGCAGAGCGTCTGGCACACTAGATTAGTAGTAAATGCTTGAT
GAGAAGAACACATCAGGCACTGCGCCACCTGCTTCACAGTACTTCCCAACAACCTTTAGAGGTAG
GTGTATTCCC GTTTTACAGATAAGGAACTGAGGCCCAGAGAGCTGAAGTACTGCACCCAGCATC
ACCAGCTAGAAAGTGGCAGAGCCAGGATTCAACCCTGGCTTGTCTAACCCAGGTTTTCTGCTCT
GTCCAATTCCAGAGCTGTCTGGTGATCACTTTATGTCTCACAGGGACCCACATCCAAACATGTAT
CTCTAATGAAATTGTGAAAGCTCCATGTTTAGAAATAAATGAAAACACCTGA

FIGURE 146

MRKHLSSWWLATVCMLLFSHLSAVQTRGIKHRIKWNRKALPSTAQITEAQVAENRPGAFIKQGRK
LDIDFGAEGNRYYEANYWQFPDGIHYNGCSEANVTKEAFVTGCINATQAANQGEFQKPDNKLHQQ
VLWRLVQELCSLKHCEFRLERGLRVTMHQPVLLCLLALIWMVK

Important features of the protein:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 157-171

N-glycosylation sites.

amino acids 98-102, 110-114

Tyrosine kinase phosphorylation site.

amino acids 76-83

N-myristoylation sites.

amino acids 71-77, 88-94, 93-99, 107-113, 154-160

Amidation site.

amino acids 62-66

FIGURE 147

GCCTTGGCCTCCCAAAGGGCTGGGATTATAGGCGTGACCACCATGTCTGGTCCAGAGTCTCATT
CCTGATGATTTATAGACTCAAAGAAAACATGTTTCAGAAGCTCTCTCTCTTCTGGCCTCCTCT
CTGTCTTCTTTCCCTCTTTCTTCTTATTTTAATTAGTAGCATCTACTCAGAGTCATGCAAGCTGG
AAATCTTTCATTTTGCTTGTCAGTGGGGTAGGTCAGTCTTAGTTTTATTTTTTGAATTT
CAACTTTCAGATTCAGGGGGTACATGTGAAGGTTTGTGTTTATGAGTATATTGCATTGATGCTGAGG
TTTGGGGT

10063549

FIGURE 148

MFRSSLLFWPPLCLLSLFLLLISSIYSECKLEIFHFACQWGRSLSLSFYFLKFQLSDSGGTCE
GLFYEYIA

Important features of the protein:

Signal peptide:

amino acids 1-25

N-myristoylation site.

amino acids 62-68

FIGURE 149

GTCTCCGCGTCACAGGAAGTTTACAGACCCACAGGGCGGACAGCGCTCCCCTCTACCTGGAGACTTGAC
TCCGCGCGCGCCCAACCTGCTTATCCCTTGACCGTCGAGTGTGAGAGATCCTGCAGCGCGCCAGTCC
CGGCCCCCTCTCCGCCCCACACCCACCCTCCTGGCTCTTCTGTTTTTACTCCTCCTTTTCATTGATA
ACAAAAGCTACAGCTCCAGGAGCCAGCGCCGGGCTGTGACCCAAGCCGAGCGTGGAAGAATGGGGTT
CCTCGGGACCGGCACTTGGATTCTGGTGTAGTGCTCCCGATTCAAGCTTTCCCAACCTGGAGGAA
GCCAAGACAAATCTCTACATAATAGAGAATTAAGTGCAGAAAGACCTTTGAATGAACAGATTGCTGAA
GCAGAAGAAGACAAGATTAAAAAACATATCCTCCAGAAAACAAGCCAGGTCAGAGCAACTATTCTTT
TGTTGATAACTTGAACCTGCTAAAGGCAATAACAGAAAAGGAAAAAATTGAGAAAGAAAGACAATCTA
TAAGAAGCTCCCCACTTGATAATAAGTTGAATGTGGAAGATGTTGATTCAACCAAGAATCGAAAAGT
ATCGATGATTATGACTCTACTAAGAGTGGATTGGATCATAAATTTCAAGATGATCCAGATGGTCTTCA
TCACTAGACGGGACTCCTTTAACCCTGAAGACATTGTCCATAAAATCGCTGCCAGGATTTATGAAG
AAAATGACAGAGCCGTGTTTGACAAGATTGTTTCTAACTACTTAATCTCGGCCTTATCACAGAAAGC
CAAGCACATACACTGGAAGATGAAGTAGCAGAGGTTTACAAAAATTAATCTCAAAGGAAGCCACAA
TTATGAGGAGGATCCCAATAAGCCACAAGCTGGACTGAGAATCAGGCTGGAAAAATACCAGAGAAAG
TGACTCCAATGGCAGCAATTCAGATGGTCTTGCTAAGGGAGAAAACGATGAAACAGTATCTAACACA
TTAACCTTGACAAATGGCTTGGAAGGAGAACTAAAACCTACAGTGAAGACAACCTTTGAGGAACCTCA
ATATTTCCCAATTTCTATGCGCTACTGAAAAGTATTGATTGAGAAAAGAGCAAAAGAGAAAGAAA
CACTGATTACTATCATGAAAACACTGATTGACTTTGTGAAGATGATGGTGAAATATGGAACAATATCT
CCAGAAGAAGGTGTTTCTACCTTGAAAACCTGGATGAAATGATTGCTCTTCAGACCAAAAACAAGCT
AGAAAAAATGCTACTGACAATATAAGCAAGCTTTTCCAGCACCATCAGAGAAGAGTCATGAAGAAA
CAGACAGTACCAAGGAAGAAGCAGCTAAGATGGAAAAGGAATATGGAAGCTTGAAGGATTCACAAAA
GATGATAACTCCAACCCAGGAGGAAAGACAGATGAACCCAAAGGAAAAACAGAAGCCTATTTGGAAGC
CATCAGAAAAAATATTGAATGGTTGAAGAAACATGACAAAAAGGGAAATAAAGAAGATTATGACCTTT
CAAAGATGAGAGACTTCATCAATAAACAAGCTGATGCTTATGTGGAGAAAGGCATCCTTGACAAGGAA
GAAGCCGAGGCCATCAAGCGCATTTATAGCAGCCTGTAAAAATGGCAAAAGATCCAGGAGTCTTTCAA
CTGTTTCAGAAAACATAATATAGCTTAAAACACTTCTAATTCTGTGATTAAAAATTTTTTGACCCAAGG
GTTATTAGAAAGTGCTGAATTTACAGTAGTTAACCTTTTACAAGTGGTTAAAACATAGCTTTCTTCCC
GTAAAAACTATCTGAAAGTAAAGTTGTATGTAAGCTGAAAAAAAAAAAAAAAAAAAAA

FIGURE 150

MGFLGTGTWILVLVLP IQAFPKPGGSQDKSLHNRELSAERPLNEQIAEAEEDKIKKTYPPENKPG
QSNYSFVDNLNLLKAITEKEKIEKERQSIRSSPLDNKLNVEDVDSTKNRKLIDDYDSTKSGLDHK
FQDDPDGLHQLDGTPLTAEDIVHKIAARIYEENDRAVFDKIVSKLLNLGLITESQAHTLEDEVAE
VLQKLISKEANNYEEDPNKPTSWTENQAGKIPKVTPTMAAIQDGLAKGENDETVSNTLTLTNGLE
RRTKTYSEDNFEELQYFPN FYALLKSIDSEKEAKEKETLITIMKTLIDFVKMMVKYGTISPEEGV
SYLENLDEMIALQTKNKLEKNATDNISKLFAPSEKSHEETDSTKEEAAKMEKEYGSLKDSTKDD
NSNPGGKTDEPKGKTEAYLEAIRKNIEWLKKHDKKGNKEDYDL SKMRDFINKQADAYVEKGILDK
EEAEAIKRIYSSL

N-glycosylation sites:

amino acids 68-71, 346-349, 350-353

Casein kinase II phosphorylation site:

amino acids 70-73, 82-85, 97-100, 125-128, 147-150, 188-191, 217-
220, 265-268, 289-292, 305-308, 320-323, 326-329, 362-365, 368-
341, 369-372, 382-385, 386-389, 387-390

N-myristoylation sites:

amino acids 143-148, 239-244

FIGURE 151

CGGCTCGAGGCTCCCGCCAGGAGAAAGGAACATTCTGAGGGGAGTCTACACCCTGTGGAGCTCAA
GATGGTCTCTGAGTGGGGCGCTGTGCTTCCGAATGAAGGACTCGGCATTGAAGGTGCTTTATCTGC
ATAATAACCAGCTTCTAGCTGGAGGGCTGCATGCAGGGAAGGTCATTAAAGGTGAAGAGATCAGC
GTGGTCCCCAATCGGTGGCTGGATGCCAGCCTGTCCCCCGTCATCCTGGGTGTCCAGGGTGGAAG
CCAGTGCTGTATGTGGGGTGGGGCAGGAGCCGACTCTAACACTAGAGCCAGTGAACATCATGG
AGCTCTATCTTGGTGCCAAGGAATCCAAGAGCTTACCTTCTACCGGCGGGACATGGGGCTCACC
TCCAGCTTCGAGTCGGCTGCCTACCCGGGCTGGTTCTGTGCACGGTGCCTGAAGCCGATCAGCC
TGTCAGACTCACCCAGCTTCCCGAGAATGGTGGCTGGAATGCCCCATCACAGACTTCTACTTCC
AGCAGTGTGACT**TAGG**GCAACGTGCCCCCAGAACTCCCTGGGCAGAGCCAGCTCGGGTGAGGGGT
GAGTGGAGGAGACCCATGGCGGACAATCACTCTCTCTGCTCTCAGGACCCACGCTGACTTAG
TGGGCACCTGACCCTTTGTCTTCTGGTTCCAGTTTGGATAAATTCTGAGATTGGAGCTCAGT
CCACGGTCTCCCCACTGGATGGTGCTACTGCTGTGGAACCTTGTA AAAAACCATGTGGGGTAAA
CTGGGAATAACATGAAAAGATTTCTGTGGGGTGGGGTGGGGGAGTGGTGGGAATCATTCTGTCT
TAATGGTAAGTACAAAGTGTACCTGAGCCCCGAGGCCAACCATCCCCAGTTGAGCCTTATA
GGGTGAGTAGCTCTCCACATGAAGTCTGTCACTCACCCTGTGCAGGAGAGGGAGGTGGTCATA
GAGTCAGGGATCTATGGCCCTTGGCCAGCCCCACCCCTTCCCTTTAATCCTGCCACTGTCTATA
TGCTACCTTTCCCTATCTCTTCCCTCATCATCTTGTGTGGGCATGAGGAGGTGGTGTATGTCTAGAA
GAAATGGCTCGAGCTCAGAAGATAAAAAGATAAGTAGGGTATGCTGATCCTCTTTTAAAAACCCAA
GATACAATCAAAATCCCAGATGCTGGTCTCTATTCCCATGAAAAAGTGCTCATGACATATTGAGA
AGACCTACTTACAAAGTGGCATATATTGCAATTTATTTTAAATAAAAGATACCTATTTATATATT
TCTTTATAGAAAAAGTCTGGAAGAGTTTACTTCAATTGTAGCAATGTGAGGGTGGTGGCAGTAT
AGGTGATTTTTCTTTTAAATCTGTTAATTTATCTGTATTTCTTAATTTTCTACAATGAAGATGA
ATTCTTGTATAAAAATAAGAAAAGAAATTAATCTTGAGGTAAGCAGAGCAGACATCATCTCTGA
TTGTCTCAGCCTCCACTTCCCCAGAGTAAATTCAAATTGAATCGAGCTCTGCTGCTCTGGTTGG
TTGTAGTAGTGATCAGGAAACAGATCTCAGCAAAGCCACTGAGGAGGAGGCTGTGCTGAGTTTGT
GTGGCTGGAATCTCTGGGTAAGGAACTTAAAGAACA AAAATCATCTGGTAATCTTTCTCTAGAAG
GATCACAGCCCCCTGGGATTCCAAGGCATTGGATCCAGTCTCTAAGAAGGCTGTCTGTACTGGTTGA
ATTGTGTCCCCCTCAAATTCACATCCTTCTTGGAACTCTAGTCTGTGAGTTTATTTGGAGATAAG
GTCTCTGCAGATGTAGTTAGTTAAGACAAGGTCATGCTGGATGAAGGTAGACCTAAATTCAATAT
GACTGGTTTCTTGTATGAAAAGGAGAGGACACAGAGACAGAGGAGACGCGGGGAAGACTATGTA
AAGATGAAGGCAGAGATCGGAGTTTTCAGCCACAAGCTAAGAAACACCAAGGATTGTGGCAACC
ATCAGAAGCTTGGAGAGGCAAGAAGAATTCTTCCCTAGAGGCTTTAGAGGGATAACGGCTCTG
CTGAAACCTTAATCTCAGACTTCCAGCCTCCTGAACGAAGAAAGAATAAATTTCCGGCTGTTTTAA
GCCACCAAGGATAATTGGTTACAGCAGCTCTAGGAACTAATACAGCTGCTAAAATGATCCCTGT
CTCCTCGTGTTTACATTCTGTGTGTGTCCCCCTCCACAATGTACCAAAGTTGTCTTTGTGACCAA
TAGAATATGGCAGAAGTGATGGCATGCCACTTCCAAGATTAGGTTATAAAAAGACACTGCAGCTTC
TACTTGAGCCCTCTCTCTGTGCCACCCACCGCCCCCAATCTATCTTGGCTCACTCGCTCTGGGGG
AAGCTAGCTGCCATGCTATGAGCAGGCCTATAAAGAGACTTACGTGGTAAAAAATGAAGTCTCCT
GCCCACAGCCACATTAGTGAACCTAGAAGCAGAGACTCTGTGAGATAATCGATGTTTGTGTTTT
AAGTTGCTCAGTTTTGGTCTAACTTGTATGCAGCAATAGATAAATAATATGCAGAGAAAGAG

FIGURE 152

MVLSGALCFRMKDSALKVLYLHNNQLLAGGLHAGKVIKGEIISVVPNRWLDASLSPVILGVQGS
QCLSCGVGQEPTLTLEPVNIMELYLGAKESKSFTFYRRDMGLTSSFESAAYPGWFLCTVPEADQP
VRLTQLPENGWGNAPITDFYFQQCD

N-myristoylation sites.

amino acids 29-34, 30-35, 60-65, 63-68, 73-78, 91-96, 106-111

Interleukin-1 signature.

amino acids 111-131

Interleukin-1 proteins.

amino acids 8-29, 83-120, 95-134, 64-103

FIGURE 153

CTTCAGAACAGGTTCTCCTTCCCCAGTCACCAGTTGCTCGAGTTAGAATTGTCTGCAATGGCCGC
CCTGCAGAAATCTGTGAGCTCTTTCCTTATGGGGACCCTGGCCACCAGCTGCCTCCTTCTCTTGG
CCCTCTTGGTACAGGGAGGAGCAGCTGCGCCCATCAGCTCCCACTGCAGGCTTGACAAGTCCAAC
TTCCAGCAGCCCTATATCACCAACCGCACCTTCATGCTGGCTAAGGAGGCTAGCTTGGCTGATAA
CAACACAGACGTTTCGTCTCATTGGGGAGAACTGTTCCACGGAGTCAGTATGAGTGAGCGCTGCT
ATCTGATGAAGCAGGTGCTGAACTTCACCCTTGAAGAAGTGCTGTCCCTCAATCTGATAGGTTT
CAGCCTTATATGCAGGAGGTGGTGCCCTTCCTGGCCAGGCTCAGCAACAGGCTAAGCACATGTCA
TATTGAAGGTGATGACCTGCATATCCAGAGGAATGTGCAAAAGCTGAAGGACACAGTGAAAAAGC
TTGGAGAGAGTGGAGAGATCAAAGCAATTGGAGAACTGGATTGCTGTTTATGTCTCTGAGAAAT
GCCTGCATTGACCAGAGCAAAGCTGAAAAATGAATAACTAACCCCTTTCCTGCTAGAAATAA
CAATTAGATGCCCCAAAGCGATTTTTTTTAAACAAAAGGAAGATGGGAAGCCAACTCCATCATG
ATGGGTGGATTCCAAATGAACCCCTGCGTTAGTTACAAAGGAAACCAATGCCACTTTTGTTTATA
AGACCAGAAGGTAGACTTTCTAAGCATAGATATTTATTGATAACATTCATTGTAAGTGGTGTTT
TATACACAGAAAAAATTTATTTTTTAAATAATTGTCTTTTCCATAAAAAAGATTACTTTCCAT
TCCTTTAGGGGAAAAAACCCCTAAATAGCTTCATGTTTCCATAATCAGTACTTTATATTTATAAA
TGTATTTATTATTATTATAAGACTGCATTTTATTTATATCATTTTATTAATATGGATTTATTTAT
AGAAACATCATTCGATATTGCTACTTGAGTGTAAGGCTAATATTGATATTTATGACAATAATTAT
AGAGCTATAACATGTTTATTTGACCTCAATAAACACTTGATATCCC

FIGURE 154

MAALQKSVSSFLMGTLATSCLLLLLVQGGAAPISSHCRLDKSNFQQPYITNRTFMLAKEASL
ADNNTDVRLIGEKLFGVSMSERCYLMKQVLNFTLEEVLPQSDRFQPYMQEVVPFLARLSNRLS
TCHIEGDDLHIQRNVQKLKDTVKKLGESGEIKAIGELDLLFMSLRNACI

Important features of the protein:

Signal peptide:

amino acids 1-33

N-glycosylation sites.

amino acids 54-58, 68-72, 97-101

N-myristoylation sites.

amino acids 14-20, 82-88

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 10-21

FIGURE 155

GGCTTGCTGAAAATAAAATCAGGACTCCTAACCTGCTCCAGTCAGCCTGCTTCCACGAGGCCTGT
CAGTCAGTGCCCCGACTTGTGACTGAGTGTGCAGTGCCCAGCATGTACCAGGTCAGTGCAGAGGGC
TGCCCTGAGGGCTGTGCTGAGAGGGAGAGGAGCAGAGATGCTGCTGAGGGTGGAGGGAGGCCAAGC
TGCCAGGTTTGGGGCTGGGGGCCAAGTGGAGTGAGAACTGGGATCCCAGGGGGAGGGTGCAGAT
GAGGGAGCGACCCAGATTAGGTGAGGACAGTTCTCTCATTAGCCTTTTCCTACAGGTGGTTGCAT
TCTTGGCAATGGTCATGGGAACCCACACCTACAGCCACTGGCCCAGCTGCTGCCCCAGCAAAGGG
CAGGACACCTCTGAGGAGCTGCTGAGGTGGAGCACTGTGCCTGTGCCTCCCCTAGAGCCTGCTAG
GCCCCAACCGCCACCCAGAGTCCTGTAGGGCCAGTGAAGATGGACCCCTCAACAGCAGGGCCATCT
CCCCCTGGAGATATGAGTTGGACAGAGACTTGAACCGGCTCCCCCAGGACCTGTACCACGCCCCGT
TGCCCTGTGCCCCGCACTGCGTCAGCCTACAGACAGGCTCCACATGGACCCCGGGGCAACTCGGA
GCTGCTCTACCACAACCAGACTGTCTTCTACAGGCGGCCATGCCATGGCGAGAAGGGCACCCACA
AGGGCTACTGCCTGGAGCGCAGGCTGTACCGTGTTCCTTAGCCTGTGTGTGTGTGCGGCCCGT
GTGATGGGCTAGCCCGACCTGCTGGAGGCTGGTCCCTTTTGGGAAACCTGGAGCCAGGTGTACA
ACCACTTGCCATGAAGGGCCAGGATGCCCAGATGCTTGGCCCCTGTGAAGTGTGTCTGGAGCAG
CAGGATCCCGGGACAGGATGGGGGGCTTTGGGGAAAACCTGCACTTCTGCACATTTTGAAAAGAG
CAGCTGCTGCTTAGGGCCGCCGAAGCTGGTGTCTGTCTCTCAGGAAAGGTTTTCAAA
GTTCTGCCCATTTCTGGAGGCCACCACTCCTGTCTCTTCTCTTTTCCCATCCCCTGCTACCCTG
GCCCAGCACAGGCACTTTCTAGATATTTCCCCCTTGCTGGAGAAGAAAGAGCCCCTGGTTTTATT
TGTTTGTTTACTCATCACTCAGTGAGCATCTACTTTGGGTGCATTCTAGTGTAGTTACTAGTCTT
TTGACATGGATGATTCTGAGGAGGAAGCTGTTATTGAATGTATAGAGATTTATCCAAATAAATAT
CTTTATTTAAAAATGAAAAA

FIGURE 156

MRERPRLGEDSSLISLFLQVVAFLAMVMGHTYSHWPSCCPKSGQDTSEELLRWSTVPVPPEPA
RPNRHPESCRASEDGPLNSRAISPWRYELDRDLNRLPQDLYHARCLCPHCVSLQTGSHMDPRGNS
ELLYHNQTVFYRRPCHGEKGTKGYCLERRLYRVSLACVCVRPRVMG

Important features of the protein:

Signal peptide:

amino acids 1-32

N-glycosylation site.

amino acids 136-140

Tyrosine kinase phosphorylation site.

amino acids 127-135

N-myristoylation sites.

amino acids 44-50, 150-156

FIGURE 157

CCGGCGATGTCGCTCGTGTGCTGCTAAGCCTGGCCGCGCTGTGCAGGAGCGCCGTACCCCGAGAGCC
GACCGTTCAATGTGGCTCTGAACTGGGCCATCTCCAGAGTGGATGCTACAACATGATCTAATCC
CCGGAGACTTGAGGGACCTCCGAGTAGAACCTGTTACAACCTAGTGTGCAACAGGGGACTATTCA
ATTTTGATGAATGTAAGCTGGGTACTCCGGGCAGATGCCAGCATCCGCTTGTGTAAGGCCACCAA
GATTTGTGTGACGGGCAAAAGCAACTTCCAGTCTACAGCTGTGTGAGGTGCAATTACACAGAGG
CCTTCCAGACTCAGACCAGACCCTCTGGTGGTAAATGGACATTTTCTACATCGGCTTCCCTGTA
GAGCTGAACACAGTCTATTTTATTGGGGCCATAATATTCCTAATGCAAATATGAATGAAGATGG
CCCTTCCATGTCTGTGAATTTACCTCACCAGGCTGCCTAGACCACATAATGAAATATAAAAAAA
AGTGTGTCAAGGCCGGAAGCCTGTGGGATCCGAACATCACTGCTTGTGAAGAAGATGAGGAGACA
GTAGAAGTGAACCTTCAACCACTCCCTGGGAAACAGATACATGGCTCTTATCCAACACAGCAC
TATCATCGGGTTTTCTCAGGTGTTTGGCCACACCAGAAGAAACAAACGCGAGCTTCAGTGGTGA
TTCCAGTGACTGGGGATAGTGAAGGTGCTACGGTGCAGCTGACTCCATATTTTCTACTTGTGGC
AGCGACTGCATCCGACATAAAGGAACAGTTGTGCTCTGCCCACAAACAGGCGTCCCTTTCCCTCT
GGATAACAACAAAAGCAAGCCGGGAGGCTGGCTGCCTCTCCTCTGCTGTCTCTGCTGGTGCCA
CATGGGTGCTGGTGGCAGGGATCTATCTAATGTGGAGGCACGAAAGGATCAAGAAGACTTCCTTT
TCTACCACCACACTACTGCCCCCATTAAGGTTCTTGTGGTTTACCCATCTGAAATATGTTTCCA
TCACACAATTTGTTACTTCACTGAATTTCTTCAAAACCATTGCAGAAGTGAGGTATCCTTGAAA
AGTGGCAGAAAAAGAAAATAGCAGAGATGGGTCCAGTGCAGTGGCTTGCCACTCAAAAGAAGGCA
GCAGACAAAGTCGTCTTCTTCTTTCCAATGACGTCAACAGTGTGTGCGATGGTACCTGTGGCAA
GAGCGAGGGCAGTCCAGTGAGAACTCTCAAGACCTCTTCCCCCTTGCCCTTTAACCTTTTCTGCA
GTGATCTAAGAAGCCAGATTCTCTGCACAAATACGTGGTGGTCTACTTTAGAGAGATTGATACA
AAAGACGATTACAATGCTCTCAGTGTCTGCCCCAAGTACCACCTCATGAAGGATGCCACTGCTTT
CTGTGCAGAACTTCTCCATGTCAAGCAGCAGGTGTGAGCAGGAAAAAGATCACAAGCCTGCCACG
ATGGCTGCTGCTCCTTGTAG

FIGURE 158

MSLVLLSLAALCRSAVPREPTVQCGSETGPSPEWMLQHDLPDGLRDLRVEFVTTSVATGDYSILMNVS
WRADASIRLLKATKICVTGKSNFQSYSCVRCNYTEAFQTQTRPSGGKWTFSYIGFPVELNTVYFIGAHNIP
NANMNEDGPSMSVNFTSPGCLDHIMKYKKKCVKAGSLWDPNITACKKNEETVEVNFTTTPLGNRYMALIQH
STIIGFSQVFEPHQKKQTRASVVIPVTGDSEGATVQLTPYFPTCGSDCIRHKGTVVLCPOQTGVPFPLDNNK
SKPGGWLPLLLLSLLVATWVLVAGIYLMWRHERIKKTSFSTTTLLPPIKVLVVYPSEICFHHTICYFTEFL
QNHCRSEVILEKWQKKKIAEMGPVQWLATQKKAADKVVFLSNDVNSVCDGTCGKSEGSPSENSQDLFPLA
FNLFCSDLRSQIHLHKYVVVYFREIDTKDDYNALSVCPKYHLMKDATAFCAELLHVKKQVSAGKRSQACHD
GCCSL

Important features of the protein:

Signal peptide:

amino acids 1-14

Transmembrane domain:

amino acids 290-309

N-glycosylation sites.

amino acids 67 - 71, 103 - 107, 156 - 160, 183 - 187, 197 - 201 and 283
- 287

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 228 - 232 and 319 - 323

Casein kinase II phosphorylation sites.

amino acids 178 - 182, 402 - 406, 414 - 418 and 453 - 457

N-myristoylation site.

amino acids 116-122

Amidation site.

amino acids 488-452

FIGURE 159

AGCCACCAGCGCAACATGACAGTGAAGACCCTGCATGGCCCAGCCATGGTCAAGTACTTGCTGCT
GTCGATATTGGGGCTTGCCCTTTCTGAGTGAGGCGGCAGCTCGGAAAATCCCCAAAGTAGGACATA
CTTTTTTCCAAAAGCCTGAGAGTTGCCCCGCTGTGCCAGGAGGTAGTATGAAGCTTGACATTGGC
ATCATCAATGAAAACCAGCGCGTTTCCATGTCACGTAAACATCGAGAGCCGCTCCACCTCCCCCTG
GAATTACACTGTCACTTGGGACCCCAACCGGTACCCCTCGGAAGTTGTACAGGCCCAGTGTAGGA
ACTTGGGCTGCATCAATGCTCAAGGAAAGGAAGACATCTCCATGAATTCCGTTCCCATCCAGCAA
GAGACCCTGGTCGTCCGGAGGAAGCACCAAGGCTGCTCTGTTTCTTCCAGTTGGAGAAGGTGCT
GGTGACTGTTGGCTGCACCTGCGTCAACCCTGTCATCCACCATGTGCAGTAAGAGGTGCATATCC
ACTCAGCTGAAGAAG

FIGURE 160

MTVKTLHGPMVKYLLLSILGLAFLSEAAARKIPKVGHTFFQKPESCPPVPGGSMKLDIGIINEN
QRVSMERNIESRSTSPWNYTWTWDPNRYPSSEVVQAQCRNLGCINAQGKEDISMNSVPIQQETLVV
RRKHQGCSVSFQLEKVLVTVGCTCVTPVHHVQ

Signal sequence:

amino acids 1-30

N-glycosylation site.

amino acids 83-87

N-myristoylation sites.

amino acids 106-111, 136-141

FIGURE 161

A C A C T G G C C A A A C A A A A A C G A A A G C A C T C C C G T G C T G G A A G T A G G A G G A G A G T C A G G A C T C C C A G G
 A C A G A G A G T G C A C A A A C T A C C C A G C A C A G C C C C C T C C G C C C C C T C T G G A G G C T G A A G A G G G A T T C
 C A G C C C C T G C C A C C C A C A G A C A C G G G C T G A C T G G G G T G T C T G C C C C C T T G G G G G G G G C A G C A C
 A G G G C C T C A G G C C T G G G T G C C A C C T G G C A C C T A G A A G A T G C C T G T G C C C T G G T T C T T G T C T
 T G G C A C T G G G C C G A A G C C C A G T G G T C C T T T C T G G A G A G G C T T G T G G G C C T C A G G A C G C T A C C
 C A C T G C T C T C C G G G C C T C T C C T G C C G C C T C T G G G A C A G T G A C A T A C T C T G C C T G C C T G G G G A C A T
 C G T G C C T G C T C C G G G C C C C G T G C T G G C G C C T A C G C A C C T G C A G A C A G A C T G G T G C T G A G G T G C C
 A G A A G G A G A C C G A C T G T G A C C T C T G T C T G C G T G T G G C T G T C C A C T T G G C C G T G C A T G G G C A C T G G
 G A A G A G C C T G A A G A T G A G G A A A A G T T T G G A G G A G C A G C T G A C T C A G G G G T G G A G G A G C C T A G G A A
 T G C C T C T C C A G G C C C A A G T C G T G C T C T C C T T C C A G G C C T A C C C T A C T G C C C G C T G C G T C C T G C
 T G G A G G T G C A A G T G C C T G C T G C C C T T G T G C A G T T T G G T C A G T C T G T G G G C T C T G T G G T A T A T G A C
 T G C T T C G A G G C T G C C C T A G G G A G T G A G G T A C G A A T C T G G T C C T A T A C T C A G C C C A G G T A C G A G A A
 G G A A C T C A A C C A C A C A C A G C A G C T G C C T G C C C T G C C C T G G C T C A A C G T G T C A G C A G A T G G T G A C A
 A C G T G C A T C T G G T T C T G A A T G T C T C T G A G G A G C A G C A C T T C G G C C T C T C C C T G T A C T G G A A T C A G
 G T C C A G G G C C C C C A A A A C C C C G T G G C A C A A A A C C T G A C T G G A C C G C A G A T C A T T A C C T T G A A
 C C A C A C A G A C C T G G T T C C C T G C C T C T G T A T T C A G G T G T G G C C T C T G G A A C C T G A C T C C G T T A G G A
 C G A A C A T C T G C C C C T T C A G G G A G A C C C C C G C C A C A C C A G A A C C T C T G G C A A G C C G C C C G A C T G
 C G A C T G C T G A C C C T G C A G A G C T G G C T G C T G G A C G C A C C G T G C T C G C T G C C C G C A G A A G C G G C A C T
 G T G C T G G C G G G C T C C G G G T G G G G A C C C C T G C C A G C C A C T G G T C C C A C C G C T T T C T G G G A G A A C G
 T C A C T G T G G A C A A G G T T C T C G A G T T C C C A T T G C T G A A A G G C C A C C C T A A C C T C T G T G T T C A G G T G
 A A C A G C T C G G A G A A G C T G C A G C T G C A G A G T G C T T G T G G G C T G A C T C C C T G G G G C C C T C A A A G A
 C G A T G T G C T A C T G T T G G A G A C A C G A G G C C C C A G G A C A A C A G A T C C C T C T G T G C C T T G G A A C C C A
 G T G G C T G T A C T T C A C T A C C A G C A A A G C C T C C A C G A G G G C A G C T C G C C T T G G A G A G T A C T T A C T A
 C A A G A C C T G C A G T C A G G C C A G T G T C T G C A G C T A T G G G A C G A T G A C T T G G G A G C G C T A T G G G C C T G
 C C C C A T G G A C A A A T A C A T C C A C A A G C G C T G G G C C C T C G T G T G G C T G G C C T G C C T A C T C T T T G C C G
 C T G C G C T T T C C C T C A T C C T C C T T C T C A A A A A G G A T C A C G C G A A A G G G T G G C T G A G G C T C T T G A A A
 C A G G A C G T C C G C T C G G G G G C G G C C C A G G G G C C G C G C G C T C T G C T C C T A C T C A G C C G A T G A
 C T C G G G T T T C G A G C G C C T G G T G G G C G C C C T G G C G T C G G C C C T G T G C C A G C T G C C G C T G C C G T G G
 C C G T A G A C C T G T G G A G C C G T C G T G A A C T G A G C G C G C A G G G G C C C G T G G C T T G G T T T C A C G C G C A G
 C G G C G C C A G A C C C T G C A G A G G G C G G C G T G G T G T C T T G C T C T C T C T C C C G G T G C G G T G C G C T
 G T G C A G C G A G T G G C T A C A G G A T G G G G T G T C C G G G C C G G G G C G C A C G G C C G C A C A C G C C T T C C
 G C G C C T C G C T C A G C T G C G T G C T G C C C G A C T T C T T G C A G G G C C G G G C G C C C G G C A G C T A C G T G G G G
 G C C T G C T T C G A C A G G C T G C T C C A C C C G A C G C C G T A C C C G C C T T T T C C G A C C G T G C C C G T C T T
 C A C A C T G C C C T C C C A A C T G C C A G A C T T C C T G G G G G C C C T G C A G C A G C C T C G C G C C C C G C G T T C C G
 G G C G G C T C C A A G A G A G A G C G G A G C A A G T G T C C C G G G C C C T T C A G C C A G C C C T G G A T A G C T A C T T C
 C A T C C C C G G G G A C T C C C G C G C G G G A C G C G G G T G G G A C C A G G G G C G G G A C C T G G G G C G G G G A
 C G G G A C T T A A A T A A A G G C A G A C G C T G T T T T T C T A A A A A A

FIGURE 162

MPVPWFLLSLALGRSPVVLSELERLVGPQDATHCSPGLSCRLWDSILCLPGDIVPAPGPVLAPTHLQTELV
LRCQKETDCDLCLRVAVHLAVHGHWEPEDEEKFGGAADSGVEEPRNASLQAQVVLSTFQAYPTARCVLLEV
QVPAALVQFGQSVGSVVYDCFEAALGSEVRIWSYTQPRYEKELNHTQQLPALPWLNVNSADGDNVHLVLNV
EEQHFGLSLYWNQVQGPPKPRWHKNLTGPQIITLNHTDLVPCLCIQVWPLEPDSVRTNICPFREDPRAHQ
LWQAARLRLTLQSWLLDAPCSLPAEAALCWRAPGGDPCQPLVPPLSWENVTVDKVLEFPLLKGHPNLCVQ
VNSSEKLQEQECLWADSLGPKDDVLLLETRGPQDNRSLEPSGCTSLPSKASTRAARLGEYLLQDLQS
GQCLQLWDDDLGALWACPMKYIHKRWALVWLACLLFAAALSLILLKKDHAKGWLRLKQDVRSGAAARG
RAALLLYSADDSGFERLVGALASALCQLPLRVAVDLWSRRELSAQGPVAVFWFAQRRQTLQEGGVVLLFSP
GAVALCSEWLQDGVSGPGAHGPHDAFRASLSCVLPDFLQGRAPGSYVGACFDRLLHPDAVPALFRTVPVFT
LPSQLPDFLQALQPPRPRSGRLQERAEQVSRALQPALDSYFHPPGTPAPGRGVGPGAGPGAGDGT

Signal sequence:

amino acids 1-20

Transmembrane domain.

amino acids 453-475

N-glycosylation sites.

amino acids 118-121, 186-189, 198-201, 211-214, 238-241, 248-251,
334-337, 357-360, 391-394

Glycosaminoglycan attachment site.

amino acids 583-586

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 552-555

N-myristoylation sites.

amino acids 107-112, 152-157, 319-324, 438-443, 516-521, 612-617,
692-697, 696-701, 700-705

FIGURE 163

GGGAGGGCTCTGTGCCAGCCCCGATGAGGACGCTGCTGACCATCTTGACTGTGGGATCCCTGGCT
 GCTCAGCCCCCTGAGGACCCCTCGGATCTGCTCCAGCACGTGAAATTCAGTCCAGCAACTTTGA
 AAACATCCTGACGTGGGACAGCGGGCCAGAGGGCACCCAGACACGGTCTACAGCATCGAGTATA
 AGACGTACGGAGAGAGGGACTGGGTGGCAAAGAAGGGCTGTGACGCGATCACCCGGAAGTCCTGC
 AACCTGACGGTGGAGACGGGCAACCTCACGGAGCTCTACTATGCCAGGGTCACCGT
 GTCAGTGGGGAGGCGGTGAGCCACCAAGATGACTGACAGGTTGAGTCTCTGACGACACTAC
 CCTCAAGCCACCTGATGTGACCTGTATCTCAAAGTGAGATCGATTGAGATGATTGTTTCATCCTA
 CCCCCACGCCAATCCGTGACAGCGATGGCCACCGGCTAACCTGGAAGACATCTTCCATGACCTG
 TTCTACCACTTAGAGCTCCAGGTCAACCGCACCTACCAAATGCACCTTGGAGGGAAGCAGAGAGA
 ATATGAGTTCTTCGGCCTGACCCCTGACACAGAGTTCTTGGCACCATCATGATTTGCGTTCCCA
 CCTGGGCCAAGGAGAGTGGCCCTACATGTGCCGAGTGAAGACACTGCCAGACCGGACATGGACC
 TACTCCTTCTCCGGAGCCTTCTGTCTCTCATGGGCTTCTCGTCGAGTACTCTGCTACCTGAG
 CTACAGATATGTACCAAGCCGCTGCACCTCCCAACTCCCTGAACGTCCAGCGAGTCTGACTT
 TCCAGCCGCTGCGCTTCATCCAGGAGCACGTCCTGATCCCTGTCTTTGACCTCAGCGGCCCCAGC
 AGTCTGGCCCCAGCCTGTCCAGTACTCCCAGATCAGGGTGTCTGGACCCAGGGAGCCCGCAGGAGC
 TCCACAGCGGCATAGCCTGTCCGAGATCACCTACTTAGGGCAGCCAGACATCTCCATCCTCCAGC
 CCTCCAACGTGCCACCTCCCCAGATCCTCTCCCCACTGTCTATGCCCAAACGCTGCCCTGAG
 GTCGGGCCCCCATCCTATGCACCTCAGGTGACCCCCGAAGCTCAATTCCCATTTCTACGCCCCACA
 GGCCATCTCTAAGGTCCAGCCTTCTCCTATGCCCCCTCAAGCCACTCCGGACAGCTGGCCTCCCT
 CCTATGGGGTATGCATGGAAGTTCTGGCAAAGACTCCCCACTGGGACACTTTCTAGTCCTAAA
 CACCTTAGGCCTAAAGGTGAGCTTCAGAAAGAGCCACCAGCTGGAAGCTGCATGTTAGGTGGCCT
 TTCTCTGCAGGAGGTGACCTCCTTGGCTATGGAGGAATCCCAAGAAGCAAAATCATTGCACCAGC
 CCTGGGGATTGACAGACAGAACATCTGACCCAAATGTGCTACACAGTGGGGAGGAAGGGACA
 CCACAGTACCTAAAGGGGCCAGCTCCCCCTCCTCTCCTCAGTCCAGATCGAGGGCCACCCCATGTC
 CCTCCCTTTGCAACCTCCTTCCGGTCCATGTTCCCCCTCGGACCAAGGTCCAAGTCCCTGGGGCC
 TGCTGGAGTCCCTTGTGTGTCCCAAGGATGAAGCCAAGAGCCAGCCCTGAGACCTCAGACCTG
 GAGCAGCCACAGAACTGGATTCTCTTTTCAGAGGCCTGGCCCTGACTGTGAGTGGGAGTCCCTG
 AGGGGAATGGGAAAGGCTTGGTGCTTCTCCTGTCCCTACCCAGTGTACATCCTTGGCTGTCA
 ATCCCATGCCTGCCCATGCCACACACTCTGCGATCTGGCCTCAGACGGGTGCCCTTGAGAGAAGC
 AGAGGGAGTGGCATGCAGGGCCCTGCCATGGGTGCGCTCCTCACCGGAACAAAGCAGCATGATA
 AGGACTGCAGCGGGGAGCTCTGGGGAGCAGCTTGTGTAGACAAGCGCGTGCTCGCTGAGCCCTG
 CAAGGCAGAAATGACAGTGCAAGGAGGAAATGCAGGGAACTCCCGAGGTCCAGAGCCCCACCTC
 CTAACACCATGGATTCAAAGTGCTCAGGGAATTTGCTCTCCTTGCCCCATTCTTGCCAGTTTC
 ACAATCTAGCTCGACAGAGCATGAGGCCCTGCCTCTTCTGTGATTGTTCAAAGGTGGGAAGAGA
 GCCTGGAAAAGAACAGGCCTGGAAAAGAACAGAAGGAGGCTGGGCAGAACCAGAACACCTGC
 ACTTCTGCCAAGGCCAGGGCCAGCAGGACGGCAGGACTCTAGGGAGGGGTGTGGCCTGCAGCTCA
 TTCCAGCCAGGGCAACTGCCTGACGTTGCACGATTTACAGCTTCATTCTCTGATAGAACAAAGC
 GAAATGCAGGTCCACCAGGGAGGGAGACACACAAGCCTTTTCTGCAGGCAGGAGTTTCAGACCT
 ATCCTGAGAATGGGGTTTGAAGGAAGGTGAGGGCTGTGGCCCTGGACGGGTACAATAACACAC
 TGTACTGATGTCACAACCTTTGCAAGCTCTGCCTTGGGTTGAGCCATCTGGGCTCAAATTCAGC
 CTCACCACTCACAAGCTGTGTGACTTCAAACAAATGAAATCAGTGCCAGAACCTCGGTTTCTC
 ATCTGTAATGTGGGGATCATAACACCTACCTCATGGAGTTGTGGTGAAGATGAAATGAAGTCATG
 TCTTTAAAGTGCTTAATAGTGCTTGGTACATGGGCAGTGCCCAATAAACGGTAGCTATTAAAAA
 AAAAAAA

FIGURE 164

MRTLLTILTVGSLAAHAPEDPSDLLQHVKFQSSNFENILTWDSGPEGTPDTVYSIEYKTYGERDW
VAKKGCQRITRKSCNLTVETGNLTelyyArvtAVSAGGRSATKMTDRFSSLQHTTLKPPDVTCIS
KVRSIQMIVHPTPTPIRAGDGHRLTLEDIFHDLFYHLELQVNRTYQMHLGGKQREYEFFGLTPDT
EFLGTIMICVPTWAKESAPYMCrvKTLpDRTWTYSFSGAFLFSMGFLVAVLCYLSYRYVTKPPAP
PNSLNVQRVLTFQPLRFIQEHVLIPVFDLSGPSSLAQPVQYSQIRVSGPREPAGAPQRHSLSEIT
YLGQPDISILQPSNVPPPQILSPLSYAPNAAPEVGPPSYAPQVTPEAQFFPYAPQAISKVQPSSY
APQATPDSWPPSYGVCMEGSGKDSPTGTLSSPKHLRPKGQLQKEPPAGSCMLGGLSLQEVTSLAM
EESQEAKSLHQPLGICTDRTSdpNVLHSGEEGTPQYLKGQLPLLSSVQIEGHPMSLPLQPPSGPC
SPSDQGPSPWGGLLESIVCPKDEAKSPAPETSDLEQPTELDSLFRGLALTQWES

Signal sequence.

amino acids 1-17

Transmembrane domain.

amino acids 233-250

N-glycosylation sites.

amino acids 80-83, 87-90, 172-175

N-myristoylation sites.

amino acids 11-16, 47-52, 102-107, 531-536, 565-570

FIGURE 165

TGGCCTACTGGAAAAAAAAAAAAAAAAAAGTCACCCGGGCCCGCGGTGGCCACAACATGG
CTGCGGCGCCGGGGCTGCTCTTCTGGCTGTTCTGCTGGGGGCGCTCTGGTGGGTCCCGGGCCAG
TCGGATCTCAGCCACGGACGGCGTTTCTCGGACCTCAAAGTGTGCGGGGACGAAGAGTGCAGCAT
GTTAATGTACCGTGGGAAAGCTCTTGAAGACTTCACGGGCCCTGATTGTCGTTTTGTGAATTTTA
AAAAAGGTGACGATGTATATGTCTACTACAACTGGCAGGGGGATCCCTTGAACTTGGGCTGGA
AGTGTGTAACACAGTTTTGGATATTTCCAAAAGATTTGATCAAGGTACTTCATAAATACACGGA
AGAAGAGCTACATATTCAGCAGATGAGACAGACTTTGTCTGCTTTGAAGGAGGAAGAGATGATT
TTAATAGTTATAATGTAGAAGAGCTTTTAGGATCTTTGGAAGTGGAGGACTCTGTACCTGAAGAG
TCGAAGAAAGCTGAAGAAGTTTCTCAGCACAGAGAGAAATCTCCTGAGGAGTCTCGGGGGCGTGA
ACTTGACCCTGTGCCTGAGCCCGAGGCATTCTCAGAGCTGATTCTCAGAGGATGGAGAAGGTGCTTTCT
CAGAGAGCACCGAGGGGCTGCAGGGACAGCCCTCAGCTCAGGAGAGCCACCCTCACACCAGCGGT
CCTGCGGCTAACGCTCAGGGAGTGCAGTCTTCGTTGGACACTTTTGAAGAAATTCTGCACGATAA
ATTGAAAGTGCCGGGAAGCGAAAGCAGAACTGGCAATAGTTCTCCTGCCTCGGTGGAGCGGGAGA
AGACAGATGCTTACAAAGTCCTGAAAACAGAAATGAGTCAGAGAGGAAGTGGACAGTGCCTTATT
CATTACAGCAAAGGATTTTCGTTGGCATCAAAATCTAAGTTTGTGTTTACAAAGATTGTTTTTAGTA
CTAAGCTGCCTTGGCAGTTTGCAATTTTTGAGCCAAACAAAATATATATTTTCCCTTCTAAGTA
AAAAAAAAAAAAAAAAAAAA

FIGURE 166

MAAAPGLLFWLFWLGALWWVPGQSDLSHGRRFSDLKVCGDEEC SMLMYRGKALEDFTGPDCRFVN
FKKGDDVYVYYKLAGGSLELWAGSVEHSFGYFPKDLIKVLHKYTEEELHIPAETDFVCFEGGRD
DFNSYNVEELLGSLELEDSVPEESKKAEEVSQHREKSPEESRGRELDPVPEPEAFRAEDSEGEA
FSESTEGLOGQPSAQESH PHTSGPAANAQGVQSSLDTFEEILHDKLKVPGSESRTGNSSPASVER
EKTDAYKVLKTEMSQRGSGQCVIHYSKGFRWHQNL SLFYKDCF

Important features of the protein:

Signal peptide:

amino acids 1-22

N-glycosylation site.

amino acids 294-298

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 30-34

Tyrosine kinase phosphorylation site.

amino acids 67-76

N-myristoylation sites.

amino acids 205-211, 225-231, 277-283

Amidation site.

amino acids 28-32

FIGURE 167

CCAGGACCAGGGCGCACC GGCTCAGCCTCTCACTTGTCAGAGGCCGGGGAAGAGAAGCAAAGCGC
 AACGGTGTGGTCCAAGCCGGGGCTTCTGCTTCGCCTCTAGGACATACACGGGACCCCTAACTTC
 AGTCCCCCAAACGCGCACCTCGAAGTCTTGAATCCAGCCCCGCACATCCACGCGCGGCACAGG
 CGCGGCAGGCGGCAGGTCCCGGCCGAGGCGATGCGCGCAGGGGGTCGGGCAGCTGGGCTCGGGC
 GCGGGGAGTAGGGCCCGCAGGGAGGCAGGGAGGCTGCATATTCAGAGTCGCGGGCTGCGCCCTG
 GGCAGAGGCCGCCCTCGCTCCACGCAACACCTGCTGCTGCCACCGCGCCGCGATGAGCCGCGTGG
 TCTCGCTGCTGCTGGGCGCGCGCTGCTCTGCGGCCACGGAGCCTTCTGCCGCCGCTGGTCAGC
 GGCCAAAAGGTGTGTTTTGCTGACTTCAAGCATCCCTGCTACAAAATGGCTACTTCCATGAAC
 GTCCAGCCGAGTGAGCTTTCAGGAGGCACGCCTGGCTTGTGAGAGTGAGGGAGGAGTCTCTCTCA
 GCCTTGAGAATGAAGCAGAACAGAAGTTAATAGAGAGCATGTTGCAAAACCTGACAAAACCCGGG
 ACAGGGATTTCTGATGGTGATTTCTGGATAGGGCTTTGGAGGAATGGAGATGGGCAAAACATCTGG
 TGCCTGCCAGATCTCTACCAGTGGTCTGATGGAAGCAATTCACAGTACCGAAACTGGTACACAG
 ATGAACCTTCTCGGGAAGTGAAGAGTGTGTGTGATGTATCACCACCAACTGCCAATCCTGGC
 CTTGGGGGTCCCTACCTTTACCAGTGAATGATGACAGGTGAACATGAAGCACAATTATATTTG
 CAAGTATGAACCAGAGATTAATCCAACAGCCCTGTAGAAAAGCCTTATCTTACAAATCAACCAG
 GAGACACCCATCAGAATGTGGTTGTTACTGAAGCAGGTATAATTCCCAATCTAATTTATGTGT
 ATACCAACAATACCCCTGCTCTTACTGATACTGGTTGCTTTTGAACCTGTGTTTCCAGATGCT
 GCATAAAAGTAAAGGAAGAACAACAACTAGTCCAAACCAGTCTACACTGTGGATTTCAAAGAGTA
 CCAGAAAAGAAAGTGGCATGGAAGTATAATAACTCATTGACTTGGTTCCAGAATTTTGAATCT
 GGATCTGTATAAGGAATGGCATCAGAACAATAGCTTGAATGGCTTGAATCACAAGGATCTGC
 AAGATGAACCTGTAAGCTCCCCCTTGAGGCAATATTAAAGTAATTTTTATATGTCTATTATTCA
 TTTAAAGAATATGCTGTGCTAATAATGGAGTGAGACATGCTTATTTTGTAAAGGATGCACCCAA
 ACTTCAAACCTCAAGCAAATGAAATGGACAATGCAGATAAAGTTGTTATCAACACGTCGGGAGTA
 TGTGTGTTAGAAAGCAATTCCTTTTATTCTTTTACCTTTTCATAAGTTGTTATCTAGTCAATGTAA
 TGTATATTGTATTGAAATTTACAGTGTGCAAAAGTATTTTACCTTTGCATAAGTGTGTTGATAAAA
 ATGAACCTGTTCTAATATTTATTTTATGGCATCTCATTTTCAATACATGCTCTTTTGATTAAAG
 AAACCTATTACTGTGTCAACTGAATTCACACACACACAAATATAGTACCATAGAAAAAGTTTGT
 TTTCTCGAAATAATTCATCTTTCAGCTTCTCTGCTTTTGGTCAATGTCTAGGAAATCTCTTCAGA
 AATAAGAAGCTATTTCAATTAAGTGTGATATAAACCTCCTCAAACATTTTACTTAGAGGCAAGGAT
 TGTCTAATTTCAATGTGCAAGACATGTGCCTTATAATTATTTTAGCTTAAATTAACAGATT
 TTGTAATAATGTAACCTTTGTTAATAGGTGCATAAAACATAATGCAGTCAATTTGAACAAAAGAAG
 TGACATACACAATATAAATCATATGTCTTCACACGTTGCCATATAATGAGAAGCAGCTCTCTGA
 GGGTTCTGAAATCAATGTGGTCCCTCTCTTGCCCACTAAACAAAGATGGTTGTTTCGGGGTTTGGG
 ATTGACACTGGAGGCAGATAGTTGCAAAGTTAGTCTAAGGTTTCCTAGCTGTATTTAGCCTCTG
 ACTATATAGTATACAAAGAGGTCATGTGGTTGAGACCAGGTGAATAGTCACTATCAGTGTGGAG
 ACAAGCACAGCACACAGACATTTTAGGAAGGAAAGGAACACGAAATCGTGTGAAATGGGTTGG
 AACCATCAGTGATCGCATATTCATTGATGAGGGTTTGCTTGAGATAGAAAATGGTGGCTCCTTT
 CTGTCTTATCTCCTAGTTTCTTCAATGCTTACGCCTTGTCTCTCAAGAGAAAGTTGTAACCT
 CTGGTCTTCATATGTCCCTGTGCTCCTTTTAACCAAATAAAGAGTTCTTGTCTTGGGGGAAAAA
 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 168

MSRVVSLLLGAALLCGHGAFCRRVVSQGKVCFADFKHPCYKMAYFHELSSRVSFQEARLACESE
GGVLLSLENEAEQKLIESMLQNLTKPGTGISDGFWIGLWRNGDGQTSGACPDLYQWSDGSNSQ
YRNWYTDEPSCGSEKCVVMYHQPTANPGLGGPYLYQWNDDRCNMKHNYICKYEPEINPTAPVEK
PYLTNQPGDTHQNVVTEAGIIPNLIYVVIPTIPLLLLILVAFGTCCFQMLHKSkgRTKTSPNQ
STLWISKSTRKESGMEV

Important features of the protein:

Signal peptide:

amino acids 1-21

Transmembrane domain:

amino acids 214-235

N-glycosylation sites.

amino acids 86-89, 255-258

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 266-269

N-myristoylation sites.

amino acids 27-32, 66-71, 91-96, 93-98, 102-107, 109-114, 140-
145, 212-217